Sudan University of Science and Technology

College of Graduate Studies

Genetic Polymorphisms of Glutathione- S-Transferase and N-Acetyltransferase-2 among Sudanese Patients with Acute Lymphoblastic Leukemia

تعدد الاشكال الجينية للجلوتاثيون اس ترانفيراس و ان استيل ترانسفيراس- 2 لدي مرضى سرطان الدم الأبيض الليمفاوي الحاد السودانيين

A thesis Submitted for Fulfillment of the Requirements for PhD Degree in Medical Laboratory Science (Hematology and Immunohematology)

Submitted By

Amna Elsadig Elsafi Ali

B.Sc in Medical Laboratory ,

SUST (2004)

M.Sc in Medical Laboratory Science Hematology and immunohematology,

SUST (2011)

Supervisored by:

Professor**:** Babiker Ahmed Mohamed

CO- Supervisored by:

Professor**:** Nazik Elmalaika Obaid Seid Ahmed

Dr**:** Mai Abdul Rahman Masri

2020

DECLARATION

I declare that this thesis is submitted to Sudan University of Science and Technology, College of Graduate studies for the degree of doctor of philosophy and has not been previously submitted by me for this degree or other degree at this university or any other universities or institutes.

Candidate Name Signature

Amna Elsadig Elsafi

Signature

الآية

بسم الله الرحمن الرحيم

قال تعالي: (وَفِي أَنفُسِكُم أَفلاَ تُبْصِرُونَ)

صدق الله العظيم

(سورة الذاريات ا لاية "21 ")

DEDICATION

To my beloved and blessed parents

Fatima and Elsadig who spared no efforts to enlighten our life

To my daughters

Laden and Shaden

To my

Sisters and friends

With infinitive love

AMNA

ACKNOWLEDGMENT

Primarly my praise and thanks should be to Allah, the almighty most gracious and most merciful, who grated me the serenity, means of strength and practice to accomplish this work.

I'm deeply indebted to my supervisor Prof. Babiker Ahmad Mohammed, Prof. Nazik Elmalaika Obaid Seid Ahmed Husain and Dr. Mai Abdul Rahman Masri for their valuable help and guidance during this patience assistance and invaluable device.

My appreciation is extending to dear sisters, friends, and other members of the department of zoology Khartoum University and department of Hematology, Khartoum Oncology Hospital.

Abstract

Leukaemia represents the second type of the ten top most common cancers in Sudan, and acute lymphoblastic leukaemia (ALL) is the first malignancy of childhood cancer in Sudan. This study aimed to study the genetic plymrphism of *N-acetyltransferase-2* (*NAT2*) {TagG590A (rs1799930), Dde A803 (rs1208), BamH1 G857A (rs1801279) and Kpn1C481T (rs1799929)} and *glutathione* S-transferases (GSTT1, M1 and P1). And to investigate the haematological picture and to determine the frequency of ALL phenotypes. Has been associated with acute lymphocytic leukemia (ALL) in the different regions of the Sudan. The blood samples are collected from 300, people of whom 150 were patients arriving to Khartoum Oncology Hospital (KOH) and the remaining 150 were healthy matched volunteered controls. Were genotyped for *GSTs* and *NAT2* in the study from 2015 to 2018. Sequence technique to PCR product for 55 (NAT2/547bp) patients and controls samples and also 11(GSTP1) patients samples by Sanger Sequencing (BGI TECH SOLU TIONS (HONGKONG), CO. China). Bioinformatics has done using various publicly available soft ware's (Sift, PolyPhen-2 I Mutant 3.0, and Chimera)the effect of each SNP was predicted.

The mean age of all subjects was 13.6 years. Their sex (63.8%) males and female (36.2%). The incidence of ALL among the adolescent group (7-18 years of age) was slightly higher (41.2%) than that registered among children group (1-6 years) (32.2%) and elderly group (>18 years of age) (22.6%). The majority of ALL patients showed 'B' immune- phenotype (75.5%). Afro-Asiatic linguistic ethnic was the most popular group (72%). According to residence, the high frequencies were reported in Khartoum (22%), Kordofan (17%), AL-Jazira and Darfur (~12% each). A significant association between affected tribes and disease frequency at states (P=0.000) was reported. The mean of the complete blood count parameters of ALL patients after treatment (Hb=10±1g/dl, RBC=4±1x10^6/ml, HCT34±1%, PLT 269± 125x10^9/L) was significantly lower (p =0.00) than that in controls (Hb= 12±1g/dl, RBC=5± 5.2 10^6 / ml**,** HCT36±3%, PLT 353±8x10^9/L) and the other parameters (WBC=19±51x10^3/ml, MCV=84±12fl, MCH=28±2pg, Blast= 8±21%) was significantly higher (p=0.00) than that in control (WBC=7.6 ±2x10^3/ml, MCV =81±3fl, MCH=27±1pg, Blast=0.00%). Regarding genetic analysis, the risk of ALL in patients with *GSTT1* null genotype was not significant (OR=1.23, 95% CI=0.74 -2.0, *P*-0.44). The *GSTM1* null genotype was significantly higher in the ALL group (79, 60.3 %) compared to controls (52, 39.78%)

and increased the risk to ALL by about two folds (OR= 1.72, 95% *CI*=1.1-2.7, *p*=0.03). The frequency of GSTP1 genotypes was also significantly different between the two groups (*P*=0.000). AG, GG and AG+GG genotypes were found to increase the risk of ALL by more than three times (*P*<0.000) compared to the AA genotype. Allele frequencies were also significantly different (*P*= 0.001, OR= 1.79, CI= 1.247-2.58). In bioinformatics results for GSTP1, (Ile/Val 105; rs1695) SNP was deleterious and probably damaging by using SIFT and Polphen-2 software, respectively. The protein of GSTP1 changed from isoleucine to valine at codon 105 by the Chimera prediction software.

The NAT2-Tag G590A (rs1799930) genotype was significantly different between the two groups (p=0.01). The AA and AA+AG genotypes were found to increase the risk of ALL by about two folds (P=0.01, 0.03, respectively) compared to the GG genotype. Allele frequencies were also significantly different (P=0.00) between ALL cases and controls. The Dde A803G (rs1208) genotype was significantly different between the two groups (p=0.02). GG genotype was found to increase the risk of ALL by more than two times (p=0.02) compared to the AA genotype. Allele frequencies were also significantly different (OR=1.16, 95% *CI*=0.77-1.74, *P*=0.01).The BamH1 G857A (rs1801279) genotype was significantly different between the two groups (p=0.00). The AA and AA+AG genotypes were found to increase the risk of ALL by about three times (*P*=0.01), compared to the GG genotype. Allele frequencies were also significantly different (*P*=0.00, OR=1.9, 95% *CI*= 1.3-2.84). Regarding Kpn1 C481T (rs1799929) genotype and allele frequencies, they were insignificant different between the two groups. The additive model exposed that SNPs with a significant association with ALL were Tag G590A (rs1799930) and BamH1 G857A (rs1801279) (OR=1.5, p=0.01). Sequence analysis (Multiple Alignment) shows changes of nucleotide from C to T at Kpn1 C481T (rs1799929) SNP and from G to A at Tag G590A (rs17 99930) and BamH1G857A (rs1801279). In bioinformatics analysis, the Tag G590A (rs179 9930) and BamH1 G857A(rs1801279)ware predicted as deleterious, probably damaging and decrease protein stability using the Sift, Poly-Phen-2 and I Mutant 3.0 soft wares respectively. Project HOPE showed that the three-dimensional structure of proteins was changed by homology modelling server from arginine (R) to glutamine (Q) at position 197 (rs1799930) and 64 (rs1 80 1279). Regarding the two other SNPs, they have no deleterious effect with the same

software. The study concludes that, ALL among Sudanese children shows male predominance, with significant different males in haematological parameters than unaffected children. The more affected age group with ALL is 7-18 years, and B-phenotype represented high-frequency rate. This study indicates that GST (M1 and P1) and NAT2 (Tag (G590A), Dde (A803G) and BamH1 (G857A) mutant genotype exhibit significant association with the risk of developing ALL among Sudanese patients. The Western and Central Sudan recorded the highest rates of ALL. Most of ALL patients are from the Afro-Asiatic ethnic group.

المستخلص

يحتل سرطان الدم المرتبة الثانية من عشرة انواع السرطانات الأكثر شيوعا في السودان، ويعتبرسرطان الدم اللمفاوي الحادعند الاطفال هو سرطان الطفولة الأول بالسودان. كما الهدف الرئسي من هذه الدراسة قياس معدلاتا الدم الكامل، نوعية الظواهرالجينية ونسبتها داخل الخلايا اللمفية البيضاء (B وT)، تحديد تعدات الاشكال الجينية مثل :,rs1279 rs180 and rs1799930, rs1799929 (*NAT2*) و (*GSTs* (P1,M1 and T1 التى يحملها مرضي سرطان الدم المفوي الحاد وانواع قبائلهم وتوزيعهم على مستوى ولايات السودان.

جمعت 300 عينة دم ، 150 عينة من المرضي بسرطان الدم اللمفاوي الحاد الذين يتلقون العلاج بمركز الخرطوم للعلاج بالاشعة و150 الباقين مطوعين اشتملت طرق التحليل علي التحليل الكامل للدم وتجدبد الطفرات الجينة الي الجينات GSTs وNAT2 بستخدام مستخلص الحامض النووى (DNA extraction) في الفترة مابين 2016 الى 2018 . بالاضافة الي تحليل (NAT2/547bp)55 عينة من الحالة والمعيار و11 عينة (GSTP1) من المرضي عن طريق التسلسل الجيني عن طريق شركة {BGI TECH SOLUTIONS (HONGKONG). CO.China} كما تم استخدام بعض السوفت وير (Sift, PolyPhen-2I, Mutant 3.0, chimera and databases) للتحليل عن طريق المعلوماتية الحيوية.

أظهرت الدراسة ان ثلثي المرضى من الذكور (63.8 %) .ومتوسط الاعمار 6 .13سنة. واعلى الاعماراصابة هم مابيين 7- 18 سنة (2 .41%) مقارنة بالأطفال من 1 - 6سنة (32.2%) سنة، والبالغين 18 years of< (22.6%) (age) معظم المصابين يحملون الظاهرةالجينية للخلايا (B) (75.5%). ومن اكثر القبائل تاثرا هي ذات العرق العربى (72%). اظهرت النتائج ان ولاية الخرطوم (22%) وكردفان الكبري (17%) والجزيرة ودارفور الكبري (12%~) الاثنين معا سجلت اعلي نسبة من المصابين باليوكيميا اللمفاوية الحادة. ولها علاقة ذو اثر ايجابي مع المرضي من القبائل السودانية (p=0.00) ومتوسط معاير الدم الكامل للمصابين باليوكيميا اللمفاوية الحادة بعد استعمالهم للعلاج كالاتي قيمة متوسط خضاب الدم (10±1g/dl)، متوسط تعداد كريات الدم (4±1x10^6/µ)، متوسط تركيزخضاب الخلية (34±1%)، متوسط تعداد الصفائح الدموية (x 269±125 **)** 10^9/L لكل مرضى السرطان اللمفاوي الحاد اقل (p=0.00) بالمقارنة مع قيم متوسط المعيار. متوسط خضاب الدم (12±1g/dl)، متوسط تعداد كريات الدم (5±5.2x 10^6/µ)، متوسط تكدس الخلايا (% 36±3)، متوسط تعداد الصفائح الدموية (x10^9/L 353±8) بفرق ذو دلالة إحصائية (p=0.00)، أما متوسط حجم الخلايا (84±12fl)، متوسط خضاب الخلية (28±2pg)، متوسط نسبة الخلايا اللمفية (54±17%)، متوسط نسبة الخلايا المريضة % (8±21) ومتوسط تعداد كريات الدم البيضاء (19±51x10^3/µl) لدى المرضي اعلى قيم بالمقارنة بالمعيار، متوسط حجم الخلايا (81±3fl)، متوسط خضاب الخلية (27± 1pg) متوسط نسبة الخلايا اللمفية (9% ±39) ومتوسط نسبة الخلايا المريضة (0.00%) بفرق ذو دلالة احصائية (p=0.00). بخصوص نتايج الجينات ان جين (*GSTT* null) ليس لديه تاثير ايجابي علي المرضي - (OR= 1.23, 95% CI=0.74 2.0, *P*-0.44)، اما جين (*GSTM1* null) موجود بنسة عالية لدى المصابين (% 60.3 ,79) بالمقارنة بالمعيار (52, 39.78%) بفرق ذو دلالة احصائة (OR=1.72, 95% *CI*=1.1-2.7, *P*=0.03). وايضا جين (GSTP1) موجود بنسة عالية لدى المرضي بالمقارنة بالمعيار (*P*=0.000)، الطراز الجينى (AG, GG و

AG+GG) يذيد نسبة الاصابة بالمرض بمعدل ثلاثة مرات، بالمقارنة بالطراز الجينى (*P*=0.000) AA. ايضا تردد الاليل بين الحالة والمعيار اظهر فرق ذو دلالة احصائة (OR= 1.79, CI= 1.247-2.58 ,*P*= 0.001). عند التحليل الطفرة (GSTP1)(Ile/Val 105; rs1695) عن طريق المعلوماتية الحيوية بستخدام السوفت وير(SIFT وPolphen) اظهرت ان لها تاثير ضار علي تكوين البروتين واحدثت تغير الحامض الاحيني من ايسوليوسين الي فالين في الموقع (codon105 بستخدام (chimera prediction software). الطراز الجينى للجين (NAT2 -Tag G590A(rs1799930) اظهر فرق بين الحالة والمعيار (p=0.00 وان الطرازالجيني (AA وAA+AG) له تاثيرباصابة السرطان بمقدار مرتين (P=0.01,P=0.03) مقارنة بالطراز الجيني (GG). وان تردد الاليل بين الحالة والمعيار له فرق ذو دلالة احصائى(P=0.00).الطراز الجينى للجين (rs1208) Dde A803G اظهر فرق بين الحالة والمعيار (P=0.02), وان الطراز GG له تاثيربالاصابة بالسرطان بمقدار مرتين واكثر مقارنة بالطراز (AA). وان تردد الاليل بين الحالة والمعيار له فرق ذو دلالة احصائى (OR=1.16, 95% *CI*=0.77-1.74, P=0.01). الطراز الجينى للجين (rs1801279) BamH1G857A اظهر فرق بين الحالة والمعيار (P=0.00). وان الطراز AA و AA+AG له تاثيربالاصابة بالسرطان بمقدار مرتين واكثر (*P*=0.01) بالمقارن بالطراز (GG). وان تردد الاليل بين الحالة والمعيار له فرق ذو دلالة احصائى (OR=1.9, 95% *CI*= 1.3-2.84, P=0.01). الطراز الجينى للجين (rs1799929) Kpn1C481T وتردد الاليل بين الحالة والمعيار ليس له فرق ذو دلالة احصائى. (rs1801279) Tag G590A (rs1799930) and BamH1G857A هاتان الطفرتان لهم تاثير قوي بالاصابة بالمرض (OR=1.5, p=0.01).

اظهر التسلسل الجيني عن طريق (Multiple Alignment) تغير C الي T من Kpn1 C481T (rs1799929) SNP and من G الي A BamH1G857A, (rs17 99930). A at Tag G590A (rs1801279) عند التجليل عن طريق المعلوماتية الحيوية بستخدام Sift, PolyPhen-2 and I Mutant 3.0 (software) ان الطفرتين (rs1799930) TagG590A و(BamH1G857A rs18 01279)) لهم تاثير ضار وايضا تاثيرفى شكل البروتين المنتج و ما حصد غير ثابت. واظهر Projec HOPE تغير الحامض الاميني من arginine (R) الي glutamine (Q) علي الموقع 197و64 بالطفرتين علي التوالي. بخصوص الطفرات الاخري ليس لها تاثير يذكر.

خلصت الدراسة الى ان مرضى سرطان الدم اللمفوى معظمهم من الذكور وان الحالات لها فرق من المعيار عند تحليل الدم الكامل لهم. وان اكثر فئة مصابة هى الفئة العمرية بين 18-7 سنه و ان الطرازالظاهرى (B) هو الاكثر شيوعا. و ان الطرازالجينى للجينات (GST(M1null و P1) Tag G590A (rs17 99930), Dde {A803G(rs1208) و NAT2{BamH1G857A (rs1801279 له تاثيرفى حدوث سرطان الدم اللمفوى الحاد لدى للسودانين. يعتبر وسط وغرب السودان من اكثر المناطق انتشارا للمرض .ومن اكثر القبائل تاثرا هي القبائل ذات العرق العربى.

LIST OF CONTENTS

List of tables

List of figures

List of abbreviations

ALL	Acute Lymphocytic Leukaemia
AML	Acute Myeloid Leukaemia
B- ALL	B- Cell Acute Lymphoblastic Leukemia
cALL	common ALL
CBC	Complete Blood Count
CC	Childhood Cancer
cDNA	Complementary DNA
CLL	Chronic Lymphoid Leukaemia
cIg	Cytoplasmic Immunoglobulin
CML	Chronic Myeloid Leukaemia
CMV	Cytomegalovirus
CNS	Central Nervous System
DNA	Deoxyribonucleic Acid
EDTA	Ethylene Diamine Tetra Acetic Acid
EBV	Epstein-Barr Virus
BL	Burkitt's Lymphoma
EMF	Exposure to electromagnetic fields
FAB	French-American-British
GST	Glutathione-S- Transferases
HB	Hemoglobin
HCT	Hematocrit
HAA	Heterocyclic Amines
MCH	Mean Cell Hemoglobin
MCHC	Mean Cell Hemoglobin Concentration
MCV	Mean Cell Volume
mIg	Membrane Immunoglobulins
NAT2	N-acetyltransferasek-2
NCR	National Cancer Register
NCBI	National Center of Biotechnology Information
NHL	Non Hodgkin Lymphoma
PCR	Polymerase Chain Reaction

Ph	Philadelphia chromosome
PLTs	Platelets
SNP	Single Nucleotide Polymorphism
RBCs	Red blood cells
RFLP	Restriction Fragment Length Polymorphism
WBCs	White Blood Cells
WHO	World Health Organization
SPSS	Statistical Package of Social Sciences
T-ALL	T-Cell Acute Lymphoblastic Leukemia
TdT	Terminal Deoxynucleotidyl Transferase
KTH	Khartoum Teaching Hospital
KOH	Khartoum Oncology Hospital

Chapter one

Introduction

Chapter One

Introduction

1.1. Introduction:

Cancer is becoming a comprehensive health problem, and remains the leading cause of death worldwide (Peres *et al*., 2019, Abodlaa *et al*., 2019). The cancer incidence in the third world, including Africa, increased dramatically in recent years (Zheng *et al*., 2017, Ferlay *et al*., 2019). Leukaemia is a wide-ranging term covering a spectrum of diseases. It is a cancer of the blood or bone marrow characterized by uncontrolled growth of specific types of white blood cells (leukocytes) (Osman, 2017, Nadkarni and Hariharan, 2019). Individuals with Leukaemia usually maintain abnormally high numbers of leukocytes in their blood. This condition results in an individual's inability to maintain certain body functions, particularly a person ability to combat infection (Reehan, 2018, Kussick, 2018). The minor rates of Leukaemia reported in sub-Saharan Africa probably stand for the letdown of diagnosis or reporting (Maude and Hunger, 2018, Jemal *et al*., 2011), and at least 54,000 new ALL cases in Asia are expected to occur each year (Al-absi *et al*., 2017).

Acute lymphoblastic Leukaemia (ALL) is a form of Leukaemia (Gatt and Izraeli, 2019, Paietta, 2018). ALL is the most common malignant disease of childhood. The incidence of childhood cancer and type varysignificantly throughout the world is according to ethnicity, socio-economic status, and gender (Tishkoff *et al*., 2009, Ibrahim and Osman, 2011, Maude and Hunger, 2018, Rafieemehr *et al*., 2019, Ferrari *et al*., 2019). Cancer in Sudanese children less than 15 years old constituted about 7% of the cancer cases recorded by the National Cancer Register (NCR) in 2009(Ahmed *et al*., 2014, Hamad, 2006). Leukaemia, has significantly increased in Sudan in the current years and it is ranked as the main cause of mortality. It represented 17% of all cancers, and ALL is recorded as the first type of Leukaemias, however, cancer registration in Sudan is dependent on hospital-based cancer recoding system (Saeed *et al*., 2014, Ahmed *et al*., 2014, Hamad, 2006). ALL have common subtypes; the precursor B-lymphoblastic which is more present in childhood and precursor T-lymphoblastic Leukaemia which more frequently occurs in older children and adolescents (Chiaretti *et al*., 2014, Ohki *et al*., 2019, Teachey and Pui, 2019).Various factors have been found to induce cancer development, including endogenous (genetic) and/or exogenous causes, such as the environment, habits and life style and chance (Khan *et al*., 2017, Errahhali *et al*., 2016). Some studies investigated the

relation between some carcinogenic factors and Sudanese with ALL, like zinc, copper, manganese, cobalt and iron (ferritin) which they an essential mineral that regulate key cellular functions such as the response to oxidative stress and DNA synthesis. It has been reported that changes in serum levels of these minerals may play an important biologic role in the initiation and development of ALL (Mohamed *et al*., 2017, Abdalla, 2016, Eiman, 2018, Nori and Ibrahim, 2018). And also GSTs, Cytochrome P450 and Quinone Oxide Reducatase1 Genes reported as association genetic factor between Sudanese with ALL (Hussen, 2019). The geneticpolym-orphisms of N-acetyltransferase-2 (*NAT2*), and glutathione S-transferase (*GSTT1, M1* and *P1*) are the main genetic factors which are involved in etiology of ALL (Arias *et al*., 2014, Elhoseiny et al., 2014, Hamad, 2006, Brisson *et al*., 2015). Both represent chief groups of phase II enzymes of the detoxification and metabolism of xenobiotics and drugs (Zapletal *et al*., 2019, Singh and Ghosh, 2019) and they have strong mutagenic and carcinogenic effect (Dobritzsch *et al*., 2019).

East Africa is a strategic area to learn human genetic diversity owing to the presence of ethnically, linguistically, and geographically different populations (Dobon *et al*., 2015, Fadhla- oui-Zid *et al*., 2011). Sudan, which is an Afro-Arab country, is an ethnic mixture of about 700 tribes that speak over 400 different languages and dialects before secession and consisted of 15 states with a population density that is noticeably uneven (Saeed *et al*., 2014, Awadelkarim *et al*., 2012, Sharkey, 2008). The Sudanese tribes were classified according to the genetic classification of African languages per Greenberg's classification (1963) into four linguistic affiliations or groups; Afro-Asiatic, Nilo-Saharan, Niger-Kordofanian, and Khoisan (Greenberg and Kemmer, 1990, Boattini *et al*., 2013, Dimmendaal, 2010, Henn *et al*., 2012, Dimmendaal, 2007). To the best of our understanding, Leukaemia in the different regions of the Sudan is unknown and no distinct study has shown the epidemiology of ALL and the relation between Sudanese linguistic groups and environmental or genetic factors in general. This study is an attempt to draw awareness to the role of *NAT2* gene polymorphisms in ALL in Sudan. The study highlights a selected *NAT2* and *GST* (P1, T1, and M1) genes SNPs and predicts the underlying phenotypic variations that can be imposed by those SNPs with demographic data as well as hematological parameters among Sudanese population with ALL. Notably, our study is considered the first descriptive epidemiological study on the geographical area and

tribe's distribution of ALL-subtypes among the Sudanese population using wet laboratory and insilico approaches.

1.2 Hypothesis:

Polymorphisms in Glutathione-s-Transferase (T1, M1 andP1) and N- Acetyltransf-erase-2 [Tag (G590A), Dde (A80 3G), BamH1 (G857A) and Kpn1 (C481T)]are associated with high risk of development of (ALL).

1.3 Rational:

Data for both incidences of leukaemias and factors which may predispose populations to leukaemias in Sudan are mostly unknown. Single nucleotide polymorphisms (SNPs) are the primary source of genetic variation, the association between SNPs for genes encoding enzymes involved in the biotransformation of xenobiotics and the susceptibility to cancers have been shown in several studies (Gill *et al*., 2016, Feki-Tounsi *et al*., 2017). Therefore, the goal of the present study was to investigate the association of genetic polymorphisms of *NAT2* [Tag G590A (rs1799 930), Dde A803G (rs1208), BamH1 G857A (rs1801279) and Kpn1C4 81T (rs1799929)] and *GST* (P1, T1, and M1) with demographic data and hematological parameter among Sudanese tribes with ALL. An understanding of the risk factors linked to increased cancer prevalence can be used to guide national health screening strategies and the development of more effective biologically- based therapies. This study will examine the hypotheses that there are different SNPs in *NAT2* and *GSTs* genes in Sudanese population from different ethnic groups, an important risk factor of leukaemia. Therefore, a new type of ALL polymorphisms is expected to emerge. Public awareness in leukaemia as a public health problem will also increase. Published studies in ALL leukaemias and myeloproliferative disorders observed a significant relation between combined *NAT2* and *GSTs* polymorphisms and development of hematopoietic neoplasms (Trang *et al*., 2018, Yu *et al*., 2016, Lukas *et al*., 2017). Globally, ALL is a blood abnormality and the major childhood cancer (Shah *et al*., 2019).

In Europe, Asia, the Middle East, and the Americas, some studies have reported that childhood ALL genetics consortium includes 12 research groups, but there is no matched study in Sudan, or Arabic world has been published except only one from Egypt **(**Mahmoud *et al*., 2010). In 2013 the proportion of new cases from ten top

cancers in Sudan, for CML was 6.4%, Lymphoma was 5.5%, and Non Hodgkin Lymphoma (NHL) was 6.9%, and there are 515 new cases of ALL among three consecutive years (2010-2012), according to the KOH records. Studies of this type are essential to the understanding of the relationship between environmental and genetic variability at multiple loci. This will improve knowledge about Leukaemia etiology and the identification of persons at increased risk of developing Leukaemia. Understanding the risk factors linked to increased cancer prevalence can inform national health screening strategies and the development of biologically based therapies through the detection of associations between the polymorphisms studied and leukemic disease risk or prognosis in Sudanese populations. The results may also have implications for pre-marital genetic counseling.

1.4 Objective

1.4.1 General objective:

To study the Genetic Polymorphisms of Glutathione-s-Transferase and N-Acetyltransferase-2 among Sudanese patients with Acute Lymphoblastic Leukemia

.

1.4.2 Specific objectives:

1. To measure the complete blood count (CBC) and relate them to demographic data (Age, Sex, geographic area, ethnic group) in ALL patients and controls.
2. To detect the frequency in [(Tag G590A (rs1799930), [Dde A803G (rs1208), BamH1 G857A (rs1801279) and Kpn1 C481T (rs179 9929)] and Glutathione-S-transferases (P1, T1 and M1) genes and alleles in Sudanese patients with ALL and control cases.
3. To study the association between the demographic data and polymorphisms in *NAT2* and *GST* (P1, T1, and M1) polymorphisms in Sudanese patients (ALL patients and control cases).
4. To detect whether polymorphic variation in *NAT2* and *GST* (P1, T1, and M1) confer susceptibility to ALL among the Sudanese population.
5. To study the present of other SNPs at (*GSTP1* and *NAT2***)** DNA sequences by multiple alignment.

Chapter two

Literature Review

Chapter two

Literature Review

2.1. Background:

The incidence of childhood cancer (CC) and type differ greatly throughout the world**.**CC contains a variety of malignant tumors. Yearly incidence worldwide is around 160.000 cases per year, whereas mortality rates average 90.000. Leukaemia is diagnosed in about 30− 34% of all CC (Inskip *et al*., 2019). The lowest rates of Leukaemia reported in sub-Saharan Africa possibly represent failure of diagnosis or reporting to some extent(Ibrahim and Osman, 2011, Erdmann *et al*., 2019).ALL is a common malignancy in children, where it corresponds to about 80% of all childhood Leukaemia's and 25-30%% of all childhood cancers. There is a slight male to female predominance of 1.3:1.0 (male to female ratio) (Zheng *et al*., 2017, Hunger and Mullighan, 2015). A majority of cases occur in persons under the age of 15 years with incidence peaks 2–5 years of age, except in Africa and the Middle East (Al-absi *et al*., 2017, Pui, 2011).

High rates of ALL, up to 40 per million, occurs in the mostly white populations of North America, Western Europe and Oceania, and also in the Chinese, and lowest of all in much of sub-Saharan Africa. In general, incidence is associated with levels of socioeconomic development (Bouda *et al*., 2020). It results from the accumulation of genetic alterations of B or T lymphoid precursor cells with arrested maturation (Churchman *et al*., 2016), and distinguished by excess growth of the lymphoblasts, which are immature lymphocytes that are activated by an antigen and is a common source of variance in a genome, and play a major role in the understanding of the genetic basis of many complex human diseases (Chiaretti *et al*., 2014). There are more than 10 million SNPs in public databases. A SNP is a single base pair change (substitution/ deletion/ insertion of one nucleotide) at a specific locus, generally consisting of two alleles. To date most association studies of ALL have been based on the candidate gene approach and have evaluated a restricted number of SNPs(Eissa and Ahmed, 2013). Such studies have highlighted difficulties in conducting statistically and methodologically rigorous investigations into ALL risk (Mathoulin-Pélissier and Pritchard-Jones, 2019, Sulong, 2010). Living organisms are always exposed to non-nutritional distant chemical agents. These xenobiotics may interact deleteriously with an organism, causing toxic and sometimes carcinogenic effects (Errahhali *et al*., 2016).

2. 2 Clinical presentation of Acute Lymphocytic Leukaemia:

The most common symptoms of ALL include fever (caused by Leukaemia or secondary to severe infections if neutropeniais present), weakness, growth retardation, bone or particular pain, petechiae and ecchymoses (Yabluchansky *et al*., 2016). More severe clinical manifestations include dyspnea, hepato and spleenomegaly, lymph adenopathy, mediastinal and testicular infiltration. ALL currents with varying levels of neutropenia, anemia, bleeding and thrombocytopenia. Leukocytosis is mainly common but symptomatic leukostasis is rare in ALL. However, a high WBC count is an independent predictor of poor effect. Central nervous system (CNS) involvement by ALL occurs in about 5–10% of adult patients through the way of their disease, but neurologic symptoms are uncommon at diagnosis (Mehta, 2011).

2.3 Classification of Acute Lymphocytic Leukaemia:

Leukaemia can be classified in a number of ways. A best classification is one which knows real entities with essential biological differences.There are four main types of Leukaemia:

- Acute Lymphoid Leukaemia (ALL)
- Acute Myeloid Leukaemia (AML)
- Chronic Lymphoid Leukaemia (CLL)
- Chronic Myeloid Leukaemia (CML)

In adults, the main types of Leukaemia are AML,CML, and CLL, but ALL is common in children(Mehta, 2011, Bashasha *et al*., 2017).

The French-American-British (FAB) Group classified ALL based on the morphology of leukemic cells and cytochemical staining of blasts. However, the latest classification system proposed by the World Health Organization (WHO) needs an additional assessment of the leukemic blasts by molecular analysis and flow cytometry (immunophenotyping). The results of these four techniques classified ALL into:

- L1, ALL with fairly small uniform lymphoblasts
- L2, ALL with more pleomorphic Lymphoblasts
- L3, ALL with basophilic vacuolated lymphoblasts

According to immunophenotyping, the ALL have two common subtypes, the precursor B-lymphoblastic which occurs more in children and precursor T-lymphoblastic Leukaemia which is more frequently present in older children and

adolescents (Wenzinger *et al*., 2018, Vadillo *et al*., 2018, DiGiuseppe, 2016). T cell–ALL accounts for about 15% and 25% of ALL in children and adults, respectively, and B cell–ALL accounts for approximately70-80% in adults (Tasian and Hunger, 2017).

2.3.1 B-lineage Acute Lymphocytic Leukaemia:

2.3.1.1 Pro-B ALL, or early pre-B, is CD10 negative and is lacking, in particular, B- or T-cell differentiation markers but expresses human leukocyte antigen (HLA)-DR, TdT and CD19, and has rearranged immunoglobulin genes. It presents in about 9–11% of adult cases.

2.3.1.2 Common ALL(c-ALL) is the main immunological subtype in adult ALL. It approximately represents > 50% of cases of adult ALL. c-ALL is described by the presences of CD10. Blast cells do not express markers that explain relatively mature B cells for example cytoplasmic or surface membrane immunoglobulins (mIg). The blast cells are positive for CD19 and TdT.

2.3.1.3 Pre-B ALL is differentiated by the expression of cytoplasmic immunoglobulin (cIg), which is missing in common ALL, but is equal to common ALL with respect to the expression of all other cell markers.

2.3.1.4 Mature B-cell ALL is found in a roundly 5% of adult cases. The blast cells express surface antigens of mature B cells, including surface mIg. CD10 may be present, in addition to, occasionally cIg (Paietta, 2018, Birva *et al*., 2019).

2.3.2 T-Acute Lymphocytic Leukaemia Classification:

About 22% of adult patients have blast cells with a T-cell phenotype. All cases express the T-cell antigen gp40 (CD7). An alternative of T-cell ALL blast cells expresses CD10 together with T-cell antigens. In the majority patients of T-cell ALL, one or more of the T-cell receptor genes are reorganized. These properties make it possible to differentiate T-cell ALL based on their stage of differentiation.

2.3.2.1 Early T-precursor ALL account for 6% of adult patients. It shows charact - eristic T-cell markers (cyCD3 and CD7) but no additional differentiation markers.

2.3.2.2 Thymic (cortical) T-ALL is the common frequent subtype of T-ALL (10%). It is explained particularly by the expression of CD1a. Surface CD3 may be occurring. Since this subtype is connected with a better prognosis, its recognition is of particular importance.

2.3.2.3 Mature T-ALL comprises about6%. The blast cells do not express CD1a but they are positive for surface CD3 (Mehta, 2011, Teachey and Pui, 2019).

2.4 Diagnosis of Acute Lymphocytic Leukaemia:

The morphologic recognition of lymphoblasts in the blood and bone marrow and their phenotypic explanation are of most essential to the correct diagnosis and categorization of ALL. These need careful evaluation of well-prepared peripheral blood and bone marrow aspirate smears, and phenotypic test of the blasts by cytochemical studies and by flowcytometry or immunohistochemistry with appropriate panel of surface and cytoplasmic markers. In the most of patients, the counts and cellularity are high, but in some there can be pancytopenia, and hypocellularity, which make the identification of the blasts more essential (Pui, 2011, Lichtman *et al*., 2006). Flowcytometry is a very significant tool for analysis and classification of ALL. Multiparameter flowcytometry allows one to know and distinguish individual cells in suspension by using fluorescent-labeled antibodies to cell lineage and separation associated antigens expressed by these cells. It provides a quick method- logy for the assignment of cell lineage and stage of differentiation in acute Leukaemia or blast crisis of CML (Birva *et al*., 2019).

2. 5 Biology of Philadelphia Positive ALL:

The Philadelphia chromosome [t (9, 22) (q34; q11)] *BCR-ABL1* occurs in less than 5% of children with ALL, but in about 30% of adult cases and in >50% of ALL patients over the age of 50years. It results from the equivalent translocation which makes a hybrid *BCR-ABL* gene. A 210 kD BCR-ABL fusion protein (p210) happens in an overwhelming proportion of patients with CML but in below a third of those with ALL. A shorter 190 kD BCR-ABL fusion protein is present in 50% of adult Ph (+ve) ALL patients and in 80% of childhood Ph+ ALL cases. The gene was present in around 30% of adults with B lineage ALL (Provan and Gribben, 2010).

2.6 Survival:

Long-standing survival rates for adult ALL patients have not significantly improved over the past two decades. Cure is a realistic objective, as ≥94% of children have continuous disease-free survival for five years and show cured, while 30–40% for patients 20–60 years old, worsening to < 15% for those older than 60 years and < 5% for those older than 70 years (Buffler *et al*., 2005).

2.7 Age Differences:

There is a clear peak of occurrence of ALL between two and four years old followed by falling rates during later childhood, adolescence, and early childhood (Chiaretti

and Foà, 2009). There is a bimodal pattern with a smaller secondary peak gradually tendency up further than 60 years old. The sharp peak in early childhood was first noted in the UK and USA in the 1930s, then next in Japan, China, and among children of African fall in the USA. The peak shows to be not present in many parts of the developing world. This change in occurrence with industrialization and affluence may give an answer to the leukemogenic etiological agents that give rise to ALL(Patterns, 2010).

2.8 Possible causes and risk factors:

So far, be sides the inherited factors, the World Health Organization (WHO) classifies four different groups of external agents as carcinogens which cause malignancy in children. These are physical, chemical, biological, and dietary components (e.g., cured meats). Wide data is now available on known and supposed risk factors for pediatric cancers, including but not limited to: early-life exposures to infectious agents (viruses, bacteria, protozoa, and fungi); parental, fetal, or childhood exposures to environmental toxins (pesticides, solvents, household chemicals and radiation (Rodgers *et al*., 2018, Williams, 2019).

2. 8.1 Genetic factors:

Genetic risk factors generally contain familial aggregations of genetic syndromes such as retinoblastoma, Li-Fraumeni syndrome, hereditary nonpolyposis colon cancer and others (GENES and ISRUPTION, 2018). In generally, in 5 − 15% of CC cases genetic factors are thought to affect the child to the development of cancer. Ecological and exogenous factors have much lower figures (5−10%), leaving the large majority of CC (75 − 90%) weekly understood and of unknown reasons. Because cancer is a multifactorial disease caused by genetic and environmental factors, it is often not easy to determine the critical time of exposure as during pregnancy or earlier. In addition, childhood cancers develop and manifest differently from one another due to various numbers of sources and distinctive clinic. The *NAT2* and *GST*(T1, M1, and P1) are remarkable genetic polymorphisms, which have been associated with increase an individual risk of developing ALL (Di Pietro *et al*., 2012, Ding *et al*., 2019).

2.8.1.1 Glutathione S-transferases (*GSTs*):

GST gene family encodes genes that are significant for certain life processes. It is a super family of genes, found virtually in all eukaryotes, involved in the detoxification and toxification mechanisms of dangerous agents including xenobiotics, pesticides, environmental carcinogens, and some chemotherapeutic drugs (Osman, 2017, Shepard

et al., 2011, Elhoseiny *et al*., 2014). Phase II enzymes catalyze conjugation of activated xenobiotics to an endogenous water soluble substrate, such as glutathione (GSH), UDP-glucuronic acid or glycine (Gatedee *et al*., 2007),by catalyzing the nucleophilic attack by glutathione synthase on electrophilic carbon, sulfur, or nitrogen molecules and converts to nonpolar xenobiotic substrates, thus avoiding their interaction with essential cellular proteins and nucleic acids (Josephy, 2010, Gra *et al*., 2010). The absolute and relative quantity of phase I and II enzyme activities be different between individuals and affect biological responses to xenobiotic exposure (Crocco *et al*., 2019). The six families of enzymes have been classified as alpha, mu, pi, theta, zeta, and omega. GST-mu (*GSTM1*), GST-theta (*GSTT1*), and GST-pi (*GSTP1*) have been studied most (Coughlin and Hall, 2002). The four cytosolic families of GSTs, including *GSTT1* andT2, *GSTM1*, *GSTP1*, located on chromosomes, 22q11.2.1p13.3 and chromosome 11q13, respectively (Nasr *et al*., 2015). *GSTM1* presents in human cancer tissue is characterized by two active alleles *GSTM1***A*, *GSTM1***B* and a non-functional null allele which results from the whole*GSTM1* gene deletion mutation. Unlike *GSTM1*, *GSTT1* is polymorphic and characterized by a functional (wild) allele and a non-functional (null) allele. This null allele outcome from total or partial deletion of the gene. Individuals who are carriers of such genotypes may therefore be at increased cancer risk (Dzian *et al*., 2012, Ramzy *et al*., 2011).

Genetic polymorphism in enzymes that metabolize xenobiotics has been associated with differences in cancer risk and adverse drug response. *GSTM1* is an important detoxification enzyme responsible for conjugation of electrophilic compounds to glutathione, thus stopping their reaction with DNA or proteins (Buffler *et al*., 2005, Sulong, 2010).The*GSTP*1 gene ,have a distance 2.48 kb of DNA and comprises 7 exons that encode for cytosolic GST enzyme (Elhoseiny *et al*., 2014). The first polymerphism identified was an A-G polymorphism at nucleotide 313 in exon 5which leads to an amino acid substitution of isoleucine (IIe) by valine (Val) at 105 amino acid position (Ile105Val). This substitution results in three *GSTP1* genotypes: they are (Ile/Ile) homozygous wild type, (Ile/Val) heterozygote and (Val/Val) homozygous (mutant) variant. The gene coding for GSTP1,displays polymorphism within the more coding region at codon 105 (Ile/Val 105; rs1695), play an important role in leukemogenesis, as it potentially alters protein function, moving back its detoxification ability for certain mutagens and carcinogens, which could effect in increased DNA

damage and mutation, and a greater risk of developing cancer (Dunna *et al*., 2014, Goodrich and Basu, 2012, Alaqidi and Alwash, 2019). The Codon 105 residues (GSTP1) is an active site for binding of hydrophobic electrophiles and the Ile-Val substitution affect substrate-specific catalytic activity and stability of the encoded protein and then their function (Chielle *et al*., 2016). Thus, an individual carrying 'val' variant is predictable to have lower detoxication potential and greater risk for cancer (Uddin *et al*., 2014, Pavlovic *et al*., 2019). The genetic polymorphisms of GST genes have been influenced of the enzyme activity among individuals and ethnic groups (Gra *et al*., 2008, Gatedee *et al*., 2007). Which was resulting in altered the expression or ability of enzymes to metabolization and then has been associated to increase risk for cancer mainly ALL (Maxwell and Cole, 2017, Gasic *et al*., 2018, Farasani, 2019). There are several reports on the predictive role of GSTs and ALL cussed, some of them described the relation between an increased environmental risk and lifestyle, and the others related it to genetic disorders (Sulong, 2010, Ibrahim *et al*., 2014, Brisson *et al*., 2015).

2.8.1.2 N-acetyltransferase 2 (*NAT2*):

The *NAT2*gene codes for the enzyme arylamineN-acetyltransferase-2(*NAT2*). It is involved in phase II of the detoxification and metabolization of xenobiotics and arylamine by N-acetylation of several arylamines, such as aromatic amines, a detoxification step, and O-acetylation of heterocyclic amines (HAA), resulting in bio-activation (Arias *et al*., 2014). It is a critical enzyme in clinical pharmacology, which have strong mutagenic and carcinogenic effect (Gra *et al*., 2010, Zheng *et al*., 2017).

NAT genes are found in prokaryotes and eukaryotes. In humans they are codified into two loci, as two polymorphic (isoenzymes) and functional genes: *NAT*1 (MIM# 108345) and *NAT*2 (MIM# 243400), with distinct functional roles. And a third loci, comprising a pseudo gene–*NATP*. *NAT2* is expressed mainly at a high level in the liver, but also expressed in almost all tissues of mammalian species and is encoded by a polymorphic gene presenting several nucleotide substitutions (Satoh, 2007, Di Pietro *et al*., 2012). Both genes (*NAT1*and 2) are located on chromosome 8. It has an open reading frame of 870-bp of 290 amino acids, and has different tissue distribution (Santos *et al*., 2016, Borlak and Reamon-Buettner, 2006). There is 81% similarity in the amino acid sequence, while proteins differ in only 55 amino acids. In the human population, more than 15 SNPs have been identified in the coding region of *NAT*1 and more than 25 SNPs in *NAT*2, with different effects (Di Pietro *et al*., 2012). About 65

allele variants of *NAT2* have been described in different human populations. These variants have between one and four nucleotide substitutions and the most frequent alleles consist of slow phenotype rs1801280 (I114T-341T>C), rs1799930 (R197Q-590G>A), rs1799931 (G286E-857G>A), and rs1801279 (R64Q). In addition to three others rs1041983 (282C>T), rs1799929 (481C>T), rs1208 (803A>G) which are synonymous single nucleotide polymorphisms or they have been identified as fast alleles (Borlak and Reamon-Buettner, 2006, Hein and Doll, 2012, Guaoua *et al*., 2014, Khlifi *et al*., 2013, Mittal *et al*., 2004). Complex diseases are the result of the interaction of genetic and environmental factors, and may be influenced by population ethnic diversity. Many recent studies have demonstrated the influence of the ethnic component in the genetic variation of *NAT*2 gene polymorphism (Arias *et al*., 2014, Guaoua *et al*., 2014, Mittal *et al*., 2004). The frequency of the acetylator phenotype differs depending on the ethnic group, and is divided into rapid, intermediate, and slow acetylator phenotypes. Thus, low or no *NAT2*activity has been associated with different levels of susceptibility to developing many kinds of cancers and drug toxicity (Di Pietro *et al*., 2012, Gra *et al*., 2010, Hein and Doll, 2012).

2.9 Major environmental risk factors:

2.9.1 Radiation:

Established evidence for increased risk of ALL includes gender, age, race, prenatal exposure to x-rays, therapeutic radiation, and specific genetic syndromes (Buffler *et al*., 2005).In children, genetic tendency syndromes and ionizing radiation are known risk issues. Contact to high-dose radiation (e.g., by survivors of atomic bomb blasts or nuclear reactor accidents) or a5-fold diagnostic x-rays before birth increased risk of developing ALL (Boothe *et al*., 2014). However, few studies state an increased risk of Leukaemia associated with residing in proximity to nuclear plants or occupational exposure to low-dose radiation (Hayes *et al*., 2001). But it is unclear whether exposure to electromagnetic fields (EMF) take part in the role of the effect with ALL, however, most studies, to date, have found little or no risk (Danker-Hopfe *et al*., 2018).

2.9.2 Infectious diseases:

The infectious diseases are still the main anxiety of the health care system. Sudan, as most sub-Saharan African countries, will face a predictable increase in cancer rate because of the more and more rising life expectancy, the wide distribution of cancer related infections, and the mounting pollution due to urbanization and extend of

intensive agriculture and industrialization (Parkin *et al*., 2008, Sitas *et al*., 2008). As in other sub-Saharan African countries, the cancer patterns would be expected to reproduce population demography, such as shorter life expectation and young population structure, and contact to a diversity of risk factors, related to both customary and new lifestyles (Black *et al*., 2003, Gondos *et al*., 2005). However, scanty documented about cancer in Sudan (Ikpatt *et al*., 2002, Hamad, 2006, Awadelkarim *et al*., 2008). It is now known that Cytomegalovirus (CMV) and Epstein-Barr virus (EBV) infect around90% of the population globally. CMV and EBV are the popular viruses that cause infections in pediatric leukemic patients as a result of Leukaemia linked immunosuppression status. And also a higher contact to herpes simplex virus types 1 and 2 (HSV-1 and 2) among ALL children (Loutfy *et al*., 2017). EBV is a herpes virus that infects and establishes a persistent infection in humans, it is linked with 50–70% of Burkitt's lymphoma (BL) in North Africa and South America and around 20% in Europe and North America (Crawford, 2001, Stiller, 2004). Viruses was development of a number of lymphoid malignancies, BL was the first neoplasm found to be related to EBV, with nearly 100% association in endemic areas (malaria areas), but less than 20% in low-incidence areas, and ALL coming in second (Ahmed *et al*., 2012). Some studies established that subclinical reactivation of cytomegalovirus (CMV), human herpesvirus-6 (HHV-6), and EBV were common in adult T Cell Leukaemia (ATL) patients in receipt of chemotherapy. Some reports have connected other viruses with various types of Leukaemia, including hepatitis B virus.(Tedeschi *et al*., 2006, Ogata *et al*., 2011). In the actual fact, EBV does not directly cause a specific disease of its own, as is the case of other pathogens, but it plays a role in the etiology of a number of diseases, as for example cancer, rheumatoid, and a co-factor to numerous autoimmune diseases (Jones, 2015). A causative link between EBV and Leukaemia, a disease with a distinctly high incidence in Sudan, has been widely investigated. According to the residence, the Western and Khartoum states registered wide extend of EBV in Sudan (Elawad *et al*., 2016, Ahmed *et al*., 2015).

2.10 Cancer in Sudan:

Cancer is one of the most important causes of death worldwide. It is estimated that the international cancer load will even increase in years to come and will reach 21.4 million new cases and 13.2 million deaths by 2030 (Mathers and Loncar, 2006). Referring to a projection of the WHO, cancer is considered as the second reason of

death in developing countries (10.4%), (Saeed *et al*., 2016). In Africa, about 715,000 new cases and 512,400 deaths were reported in 2008. Thus, cancer poses a massive burden for the health system, as well as the whole economy, throughout Africa because of both treatment prices and drop-out of working power of patients (Jemal *et al*., 2012). Sudan is well-explained as an Afro-Arabian country because of its rich merger of ethnic, social and cultural diversity. As of 2013, the Sudanese population is evaluated at about 36 million. Sudan's cancer trouble is not well- recorded because the awareness of the health system mainly focuses on communicable diseases such as malaria, tuberculosis and human immunodeficiency virus (HIV), which is similar to all developing countries in Africa. A review of Sudanese hospitals conducted in 2000 exposed that cancer was the third reason of death after malaria and viral pneumonia, accounting for 5% of all deaths. Most published data on cancer in Sudan are little, due to the hospital-based case studies (Osman *et al*., 2010, Ahmed *et al*., 2014). The main sources of cancer data are: KOH, located in the national capital Khartoum, Khartoum State and the National Cancer Institute (NCI) of the University of AL-Gezira in Wad Madani, Capital of the Gazira State, both centers are located in the densely populated center of Sudan. A National Cancer Registry was established in 2009 to pool cancer data into a single centrally easy to get to system. However, great and reliable datasets on cancer frequency and mortality in Sudan are still lacking (Saeed *et al*., 2016, Saeed *et al*., 2014).

2.11 Ethnicity and tribes' groups:

Sudan is a diverse country combining various ethnic, linguistic groups and cultures (Dimmendaal, 2014, Blench and MacDonald, 2006).The three out of four language families attested in Africa include Sudan per Greenberg's classification are (i.e. Afro-Asiatic, Nilo-Saharan, and Niger-Kordofanian) and Khoisan (Miller, 2016, Lewis, 2009, Henn *et al*., 2012, Tishkoff *et al*., 2007, Estey *et al*., 2014). Afro-Asiatic, spoken mostly by northern and eastern African pastoralists and agro-pastoralists, covering North Africa, includes the Semitic, Cushitic, and ancient Egyptian (Coptic) languages (Dimmendaal, 2011, Scheinfeldt *et al*., 2010). Nilo-Saharan, spoken mainly by eastern and central African pastoralists, includes in its main Chari-Nile branch the Central and Eastern Sudanic languages. Niger-Kordofanian, spoken mostly by agriculturalist populations across western, eastern, central, and southern Africa (Paul *et al*., 2009, Boattini *et al*., 2013).

Figure 2-1: Amap showing the distribution of different linguistic groups in Sudan

Chapter three

Materials and Methods

Chapter three
Materials and methods

3.1. Study design:

This is a case – control study to detect the association hematological parameters and polymorphisms of NAT2 [Tag G590A (rs1799930), Dde A803G (rs1208), BamH1 G857A (rs180 1279), and Kpn1 C481T (rs1799929)] and GST (P1, T1 and M1) genes in ALL patients and a matched healthy Sudanese population during the period from August 2015- to August 2018.

3.2. Study area:

The research was carried out at KOH, because of the facilities there and the availability of samples representative of the ethnicities throughout the country.KOH it was the first oncology in Africa established 1967, it serves Sudan and the adjacent countries.

3.3. Study population:

Study population included 150 patients attending KOH, who were diagnosed as ALL of both sexes. And 150 control patients who consisted of healthy volunteers without a medical history of cancer or other chronic diseases attending Khartoum Teaching Hospital (KTH) at the same period.

3.4. Inclusion criteria:

Patients who have confirmed diagnosis of ALL regardless of gender or ethnic group and attended KOH during the study period and the control group was collected from KTH, all patients and controls were of Sudanese ethnicity and matched for all demographic data (figure 1: see the map of the distribution of different linguistic groups in Sudan).

3.5. Exclusion criteria:

Non-Sudanese patients and those who refused to participate in the study.

3.6. Sample Size:

Sample size was determined based on preliminary data collected in a pilot survey using the following equation,

$$n = \frac{\left(\frac{Z}{d}\right)^2 \times (0.50)^2}{1 + \frac{1}{N}\left[\left(\frac{Z}{d}\right)^2 \times (0.50)^2 - 1\right]}$$

n= sample size,

z= value (e.g. 1.96 for 95% confidence level),

d = confidence interval, expressed as decimal (e.g., .05 = ±5),

N= percentage picking a choice expressed as a decimal. The cases of ALL in the previous three years were 515 cases (N), (KOH statistic office, 2014).

Accordingly, the sample size was 150 patients, and other 150healthy individuals were included as control group. The samples were collected randomly.

3.7. Blood sample collection:

Venous blood samples (3- 5 ml) were collected under sterile conditions from each case and control into two labeled EDTA vacutainer tubes: one for hematologic analysis and the second for DNA extraction.

3.8. Data collection and analysis:

Data were collected using a purposeful data collection form for 150 Sudanese patients diagnosed with ALL (Band T phenotype), treated at Khartoum Oncology Hospital (KOH) from 2015-2017 year. The form including demographic data (age, gender, tribe, and residence), family history of ALL, date of diagnosis and nutrition was used. The Statistical Package of Social Science (SPSS version 23; SPSS Inc, Chicago, IL) and GraphPad Prism version (6.07) were used for data analysis collected data were analyzed to obtain the mean, frequency, the stander deviation, Fisher exact test, probabilities (p value) and logistic regression. The 95% confidence level was used. Arc map analysis program was used to represent a map, using geographic information system (GIS). For ethnic group data using Greenberg's classification and related studies (Henn *et al*., 2012, Tishkoff *et al*., 2007).

3.9. Age groups:

Age was categorized into four groups (infant< one year, 1-6years (children),7-18years (adolescent) and>19years (old)using World Standard Population (WSP) method (Ahmad *et al*., 2001),including the 2010 Sudan population (http://www.cbs.gov.sd), and 2009 projected Khartoum state populations (Saeed *et al*., 2014).

3.10. Ethnicity and tribes' groups:

The Sudanese linguistic groups included in this study were classified to the three language families attested in Africa as referred to per Greenber- g's classification are: Afro-Asiatic, Nilo-Saharan, and Niger-Kordofanian (Henn *et al*., 2012, Miller, 2016, Estey *et al*., 2014, Dimmendaal, 2010).

3.11. Ethical consideration:

Selected individual were informed with detailed objectives of the study and its importance in the future. They were also, provided with information about the study and any risk that may arise especially during the collection, before they/their guardians signed written informed consent. All information obtained from patients was kept as highly confidential data, specimens, and results. Ethical approval was obtained from the Ministry of Health (The acceptation number: 61–1522), Khartoum State and permission has been taken from the administrative of KTH and KOH.

3.12. Laboratory methods:

3.12.1. Hematological tests:

EDTA blood samples 5 ml for hematologic analysis (Complete Blood Count and blood Film). The Hematological parameters were done using an automated particle cell counter (Sysmex KX 21N.Model 250VT2A/ Japan), and staning blood films for morphology to confirm the Sysmex result.

3.12.1.1 Sysmex KX 21N:

3.12.1.1.1 Principle:

Blood sample is aspirated, calculated to a predetermined amount, diluted at the specified ratio, and then fed into each transducer. The transducer chamber has minute hole named the aperture. On both surfaces of aperture, there are the electrodes among which flows direct current (DC). Blood cells suspended in the sample pass via the aperture, causing DC confrontation to change between the electrodes .As DC confrontation change the blood cell size is detected as electric pulses. Blood cell count

is calculated by counting the pulses and a histogram of blood cell sizes is plotted by determining the pulses sizes. Also, analyzing a histogram creates it possible to obtain a variety of analysis information (Sysmex KX 21N.Model 250VT2A/ Japan).

3.12.1.1.2 Method of complete blood count:

Completely automated multichannel devices require only that an suitable blood sample is presented to the apparatus and typically calculate from 8 – 20 components for the essential CBC and blood cell differential, impedance counting systems depends on the reality that red cells are poor conductors of electricity, while certain diluents are good conductors (Sysmex KX 21N.Model 250VT2A/ Japan).

3.12.1.2 Flowcytometry:

Immunophenotyping Flowcytometer was used to diagnose of ALL; Lymphocytes panel was used including CDs markers (Coulter EPICS XL–Mcl TM Flowcytometer –Miami, Florida –USA). Basic principle is the passage of cells in single file in front of laser beam so they can be detected, counted and sorted. Cell components are fluorescently labeled and then excited by the laser to emit light at varying wavelengths. The majority immunophenotyping studies were carried out from blood or marrow aspirates with surface and cytoplasmic markers by flow cytometry in ALL. CD34, HLA-DR, TdT, CD45 were used as general markers. CD10, CD19, cCD22, cCD79A, Igm, IgG, Kappa, Lambda, and sIg were B-cell markers used and as T-cell markers we used: CD1a, CD2, cCD3, CD4, CD8, CD5, CD7.

3.12.2. Molecular tests:

3.12.2.1 DNA isolation:

Genomic DNA was extracted from the lymphocytes donor's cell by using salting out methods with some modifications: Briefly, the pellet of White Blood Cell (WBC) was incubated in lysis buffer (10 mM Tris HCl, pH 8.0, 400 mM NaCl, 2 mM EDTA, 200 µl of 10% Sodium Dodecyl Sulfate (SDS) and 5 µl of Proteinase K) for overnight at 37C°. The solution mixture was mixed with 1ml Guanidine Chloride and 300µl potassium acetate, shaken and centrifuged at 6,000g for 5 minutes. DNA was precipitated from the resulting supernatant by adding absolute ethanol and then was washed 3 times (70% ethanol). DNA pellet was then dissolved in doubled distilled water (DD H_2O) and stored at –20 °C till used(Miller *et al*., 1999, Taspinar *et al*., 2008, Farivar and Ghazimoradi, 2019).

3.12.2.2 Polymorphism analysis:

3.12.2.2.1 Genotyping of *GSTP1* gene polymorphism:

For genotyping of the *GSTP1* (Ile105Val) polymorphism, polymerase chain reaction restriction and fragment length polymorphism (PCR-RFLP) assay was used. PCR reaction was done using one pair of primers (Table1). Amplification was carried out in 25 µl reaction mixture containing 5µl leaflet ready master mix, 16 µl double distil water, 1µl of each primers and 200 ng of genomic DNA. After an initial denaturation at 94 °C for 3min, the DNA was amplified through 35 cycles of 94°C for 30s at 56 °C for 30s and 72°C for 30s, with a final extension at 72°C for 5 min in a G-STORM PCR machine. PCR products were then electrophoresed on 2% agarose gel, to allow detection by ethidium bromide staining (Joseph *et al*., 2004). The single nucleotide substitution (A →G) at position 313 of GSTP1gene (AAT to GAT) created a BsmAI (BstMA1) recognition cleavage site for that enzyme. 5 µl of PCR products and5µl of reaction mixture (0.5U BsmA1I reaction buffer and volume was completed with ddH_2O) were incubated at 55°Cfor two hours. The products were separated using 3% agarose gels stained with ethidium bromide. The digested products were expected to produce 3different genotype patterns; the A/A homozygote wild type, demonstrated as a single 224 bp fragment, the polymorphic G/G homozygote mutant presented by 146bp and 78bp fragments, whereas the A/G heterozygote genotype exhibited 224 bp,146 bp and 78 bp fragments.

3.12.2.2.2 Multiplex (PCR) *GSTM1* and *GSTT1* Genotyping:

A multiplex PCR was used to amplify both *GSTM1* and *GSTT1* in a single PCR reaction together with ß-globin gene primer as an internal control (Table1). An initial denaturation at 95 °C for 5 minutes was followed by 35 cycles of 1 minute at94 °C, 30 seconds at 64°C, 1 minute at 72°C and final extension was for10 minutes at 72 °C. Genotyping of the genes (null genotypes) was revealed by the absence of the 480bpfor GSTT1 and 219 bp for GSTM1 PCR products. The ß-globin gene produces a 268 bp was used as an internal positive control. PCR products were visualized by2% agarose gel electrophoresis with ethidium bromide. The absence of ß-globin amplification indicated a failure of PCR reaction (Santovito *et al*., 2010).

3.12.2. 2.3 Genotyping of NAT2:

For genotyping of the *NAT2* (C481T, G590A, A803G and G857A) polymorphism, PCR-RFLP assay was used. PCR reaction was done using one pair of primers (Table1).After an initial denaturation at 95°C for 5min, the DNA was amplified through 35 cycles of 94°C for 30s at 58°C for 1min and 72°C for 30s, with a final extension at 72°C for 10 min in a G-STORM PCR machine. PCR products were digested with KpnI, TagI, DdeI and BamHI. The C481T mutant allele was identified by the loss of a KpnI restriction site and the A803G gain of a DdeI site, G590A, by the loss of a TagI site, and G857A, by the loss of a BamHI site. Restriction digests were separated by electrophoresis on 6% agarose gel to allow detection by ethidium.

3.12.3. Sequence technique:

Sanger Sequencing to PCR product for 11 patients sample (GSTP1- 176bp) and 55 (NAT2- 547bp) samples (45 patients and 10 controls) by BGI TECH SOLUTIONS HONGKONG.CO, China.

3.12.4 Bioinformatics Analysis:

Bioinformatics analysis for each sample, the two purified chromatogram (forward and reverse) nucleotide sequences were viewed and checked for quality by FinchTV program version. The NCBI Nucleotide database was searched for reference sequences. Target gene nucleotide sequence was obtained and all regions were analyzed accordingly. Additional high similarity sequences were obtained from NCBI database and were added as control sequences using nucleotide Basic Local Alignment Search Tool (BLAST). Any apparent changes within the tested sequences were noticed through multiple sequence alignment using BioEdit software. All sequences were translated into amino acid sequences using online Expasy translate tool. The resulted amino acid sequences were compared all together using BioEdit software. A computational examination of the reported GSTP1 and *NAT2* SNPs using various publicly available softwares (Sift,Polyphen, I Mutant 3.0, ProjecHOPE, chimera) and databases (NCBI, dbSNP, Expasy UniProtKB). Were conducted to understand their possible influence on the protein structure and function, and hence their ability to cause or enhance advance ALL development.

Table 3-1: Details of primers (forward and reverse) restriction enzymes and product sizes (Both before and after digestion).

Genotypic Locus	Restriction enzymes	Primers	DNA products size before digestion	DNA products size after digestion
GSTP1	BsmAI	F 5′-ACCCCAGGGCTCTATGGGAA-3′ R 5′-TGAGGGCACAAGAAGCCCCT-3′	176	91, 85
GSTT1		F5′-TCACCGGATCATGGCCAGCA-3′ R5′-TTCCTTACTGGTCCTCACATCTC-3′	480	
GSTM1		F 5′-GAACTCCCTGAAAAGCTAAAGC-3′ R 5′-GTTGGGCTCAAATATACGGTGG-3′	219	
NAT2	KpnI, TaqI, Dde, and BamHI	*F*-5′GCTGGGTCTGGAAGCTCCTC-3′, R-5′TTGGGTGATACATACACAAGGG-3′	547	To detect the mutation of C481T, C590A , A803G and G857A respectively

Key: F= Forward R= Reverse(Taspinar *et al*., 2008)

Chapter four

Results

Chapter four

Results

4.1 Demographic Data:

A number of 150 ALL patients attended at (KOH) and 150 controls were detectable for the molecular examination. The demographic information for all data was given. No different appearance between the patient and control groups, except for patients distributions within states. In 150 subjects, slightly high frequencies were registered from Kordofan (20%), Khartoum (19%), Northern (15.3%) and Darfur (13%) with significant association between ALL frequencies among affected tribes and inhabited states (P= 0.002) figure (4-2). The other demographic information including the Sudanese linguistic ethenic groups, and ALL immune phenotypes are demonstrated in Table (2), with insignificant differences in the distribution between age group (p=0.97) (Table 2).

Almost two-thirds were males (63.8%) when compared with females (36.2%, P <0.0001). The mean age of all subjects was 13.6 years. The calculated mean of age according to the different age groups was 3.9, 11.2 and 34.6 for children, adolescent, and elderly groups, respectively. The frequency among adolescent group (7-18 years of age) was slightly higher (36.9%) than that registered among children group (1-6 years) (30.7%) and elderly group (>18 years of age) (30.3%). Age-specific rates sharply decreased within the infant age group (2.1%) figure (4-3). Correlation between gender and age group figure (4-4). Afro-Asitic lingusitic group was the most common group (72%). The most affected individuals belonged to Gaalin (19%) followed by Baggara (11%), Kawahla and Hasania (~8% each) while lower representation was reported for other groups Figure (4-5). A high significant association between affected tribes and disease frequency at states (P = 0.000) (Table 2) was reported. Patient's Sudanese tribes included in study pointed out in table (3).

4.2 Characteristics of cases and controls:

4.2.1 Hematological examination results:

The analysis of complete blood count of the group of cases and controls revealed that the mean of some complete blood count parameters of ALL patients (HB=10±1 g/dl, RBC=4±1x10^6/µl, HCT=34±1%, PLT =269±125x10^9/L) was significantly lower (P=0.00) than that in controls (HB=12±1g/dl, RBC=5±5.2 x10^6/µl**,** HCT

=36±3%, PLT=353±8x10^9/L) and the mean of the other parameters (WBC= 19±51 x10^3/µl, MCV =84±12fl, MCH=28±2pg, Blast=8±21%) was significantly higher (P=0.00) than that in control (WBC= 7.6±2x10^3 /µl, MCV=81 ±3fl, MCH=27±1pg, Blast=0.00%) as seen in Table (4).

4.3. Molecular results:

4.3. 1 *GSTM1* and *GSTT1* polymorphisms:

Multiplex PCR of *GSTM1*, *GSTT1* and ß–globin as positive control showed the expected band sizes and patterns figure (4-6).Subjects with *GSTM1* null genotype were significantly higher (*P*=0.03) in the cases group (78; 52 %) compared to controls (52; 40%).The risk of ALL was almost doubled among patients with *GSTM1* null genotype (OR=1.72, 95% *CI*=1.1-2.7, *P*=0.03) (Table 5). On the other hand, the differences in the distribution of GSTT1null genotypes among cases and controls (35% and 30%, respectively) was found to be not significant (*P*=0.44) (OR=1.23., 95% *CI*=0.74-2.0, *P*=0.44) (Table 5). Double null (21%) of GSTT1and M1among cases (Table 6).

4.3.2 *GSTP1* polymorphisms:

PCR of *GSTP1* showed the expected band sizes and patterns (Figure 4-7)The distribution of theGSTP1 genotypes (AA, AG and GG genotypes; 34.7%, 44% and 21.3%*vs* 66.1%, 25.4% and 8.5% for cases and controls, respectively) was significantly different between patients and controls (*P*=0.000). AG, GG and AG+GG genotypes increased the risk of ALL by more than three times (OR=3.308, 95% *CI*=1.925-5.685, *P*=0.000; adjusted OR= 4.811, 95% *CI*= 2. 23-10.3, *P*=0. 00; adjusted OR= 3. 684, 95% *CI*= 2.24-6. 043, *P*=0.000, respectively) compared to the AA genotype. No significant association between age groups and GSTP1 genotypes. Allele frequencies were A =170 (56.6%) cases and 205(78.8%) controls, G=130 (43.3%) cases and 55 (21.1%) controls with (*P*= 0.001, OR=0.557, *CI*=0.387-0.801). Figure (4-8) shows similar multiple alignments for GSTP1 patient sequence (11 cases) with reference databases sequence using BioEdit. Bioinformatics analysis of GSTP1 the Isoleucine amino acid converted to valine at codon 105 using Chimera (1.8) Figure 4-9.

4.3.3. Regression Test:

The binary regression for risk factors for GSTp1, WBCs and HB (*P*= 0.000, OR=0.51, 95% *CI* =0,323-0.817, *P*=0.01, OR= 1.0, 95% *CI*= 1.0-1.10, *P*=0.00, OR= 2.4, 95% CI= 1.86-3.86, respectively) is illustrated in Table 7.

4.3.4. NAT2 polymorphisms (rs 1799930, rs1208, rs1801279 and rs1799929):

Tag G590A (rs1799930). The distribution of the NAT2 genotypes (GG, AG and AA genotypes; 32%, 42% and 26% *vs* 46.4%, 38.1% and 15.5% for cases and controls, respectively) was significantly different between two groups (*P*=0.01). AA and AA+AG genotypes increased the risk of ALL by more than two times (OR =2.43, 95% *CI*=1.14 -4.70, *P*=0.01 and OR=1.83, 95% *CI*=1.02- 3.22, *P*=0.03respectively) compared to the GG genotype. Alleles frequency A =94 (47%) cases and 76 (34.5%) controls, G =106 (53%) cases and 142 (64.5%) controls with OR= 1.77, 95% *CI*=1.2-2.64, *P*=00. Gel electrophoresis pattern of Tag viewed in Figure (4-10). The Dde A803G (rs1208) genotypes (AA, AG and GG) were also significantly different between cases and controls (31%, 37% and 32% *vs.* 41.8%, 40% and 18.2% for cases and controls, respectively). The GG genotype significantly increased the risk of ALL by more than two times (OR=2.37, 95% *CI*=1.15-4.88, *P*=0.02). Alleles frequencies A =99(49.5%) cases and 136(61.8%) controls, G =101(50.5%) cases and 84(38.2%) controls. OR=1.16, 95% *CI* =0.77-1.74, *P*=0.01.The Gel electrophoresis patterns of Dde are viewed in figure (4-10).

Regarding the BamH1G857A (rs1801279) genotypes (GG,AG and AA), the reported frequencies were 22%, 31% and 47% *vs.* 37.3%, 32.7% and 30% for cases and controls, respectively with significant difference between the two groups (*P*=0.00). Collectively, the AA and AA+AG genotypes increased the risk of ALL by more than two times (OR=2.65, 95% *CI*=1.34-5.23, *P*=0.00 and OR=2.107 95% *CI*=1.15 -3.88 *P*=0.01 respectively) compared to the GG genotype. Alleles frequency A =125 (62.5%) cases and 102 (46.4%) controls, G =75(37.5%) cases and 118(53.6%) controls, OR1.9, 95% CI= 1.3-2.84, *P*=0.00. Gel electrophoresis patterns of BamH1 digestion are viewed in figure (4-11).

The genotype frequencies for the Kpn1C481T (rs1799929) SNP were 25%, 45% and 30% *vs*.31.8%, 44.6% and 23.6% for the CC, TC and TT genotypes in cases and controls, respectively; The TT and TC genotypes increased the risk of ALL by more than one time (OR=1.6, 95% CI=0.77-5.53, P=0.19) compared to the CC genotype)

but did not reach significance (P=0.45). Alleles frequencies T =113(56.5%) cases and 96(43.6%) controls, C =87(43.5%) cases and 124 (56.4%) controls, OR=1.67, 95% CI=1.14-2.46, *P*=0.01.Digestion pattern of the Kpn1C481T SNP are viewed in Figure (14). The logistic regression models (controlled for age and gender) revealed that the additive genetic models of NAT2 SNPs Tag G590A (rs1799930) (OR=1.45, CI=0.995-2.136, P=0.05) and BamH1G857A (rs1801279) (OR=1.48 CI=1.048-2.106, P=0.02) were significantly associated with ALL (Table 9).

Multiple alignments for NAT2 (547bp) patient sequence forward sequence of the studied ALL patients show heterozygote nucleotide at position 92 (TC) Figure 15, also. For (NO 80-115) and control sequence (NO 121 and 122) with reference databases sequence using BioEdit Sequence shows change of wild type nucleotide (C) to mutant (T) for Kpn1C481T(rs1799929) SNP (Figure 16), and wild type nucleotide (G) to mutant (A) for Tag G590A (rs1799930) and BamH1G857A (rs180 1279) Figure(4-15).

4.4. NAT2 Bioinformatics results:

The rs1799930 (R197Q) and rs1801279 (R64W) SNPs are expressed as deleterious (score=0.00), probably damaging (score=1.00) and decrease protein stability SNPs (rs 1799930 =-1.26DDG and rs1801279= -0.05 DDG) with Sift, PolyPhen-2 and I-Mutant 3.0 respectively. And their three-dimensional structure of proteins changed by ProjecHOPE homology modeling server, from arginine (R) to glutamine (Q) at position 197 (rs1799930) (figure 4-16). And arginine (R) to tryptophan at 64 (rs180 1279) by Chimera (Figure 4-17).

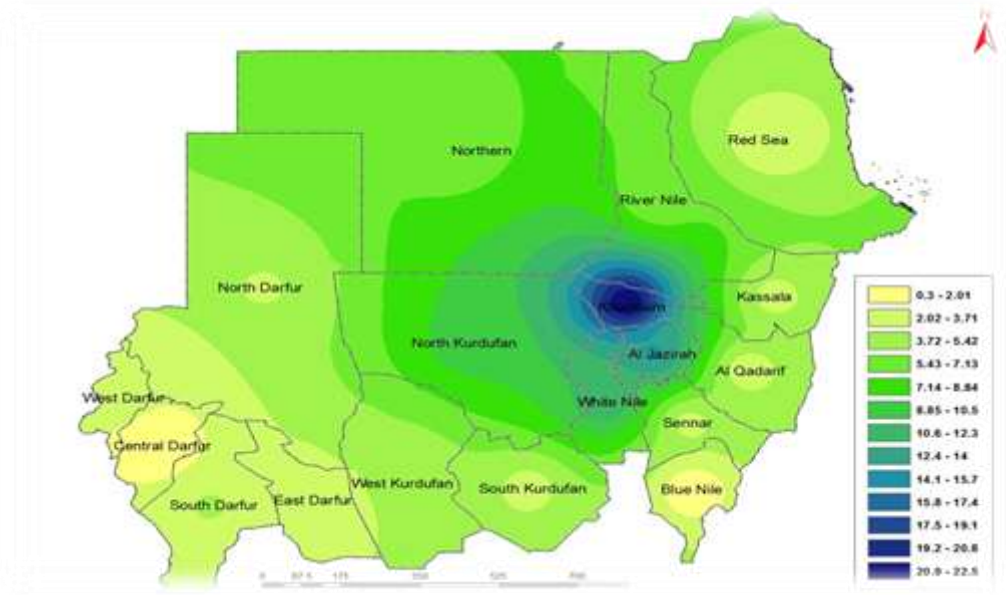


Figure 4-2: A map of states of Sudan, showing the distribution of ALL cases attended Khartoum Oncology Hospital during 2015-2017.

Table 4-2: Demographic data and ALL -subtype distribution in the study population.

Characteristic	Total	P .Value
Gender:		
Male: Count (%)	568(63.8%)	0.000
Female: Count (%)	322(36.2%)	
Age:		
1-< one year; Count (%)	19 (2.1%)	0.000
2-(1-6 year) ;count (%)	273(30.7%)	
3-(7-18 year) ;count (%)	328(36.9%)	
4-> 18year; count (%)	270(30.3%)	
Tribe:		
1-Afro-Asiatic count(%)	641 (72%)	0.000
2-Nilo-Saharan count(%)	156(17.5%)	
3- Niger-Kordofanin Count(%)	93(10.4%)	
Immunophenotypes:		
B-ALL	668(75.1%)	0.000
T-ALL	222(24.9%)	

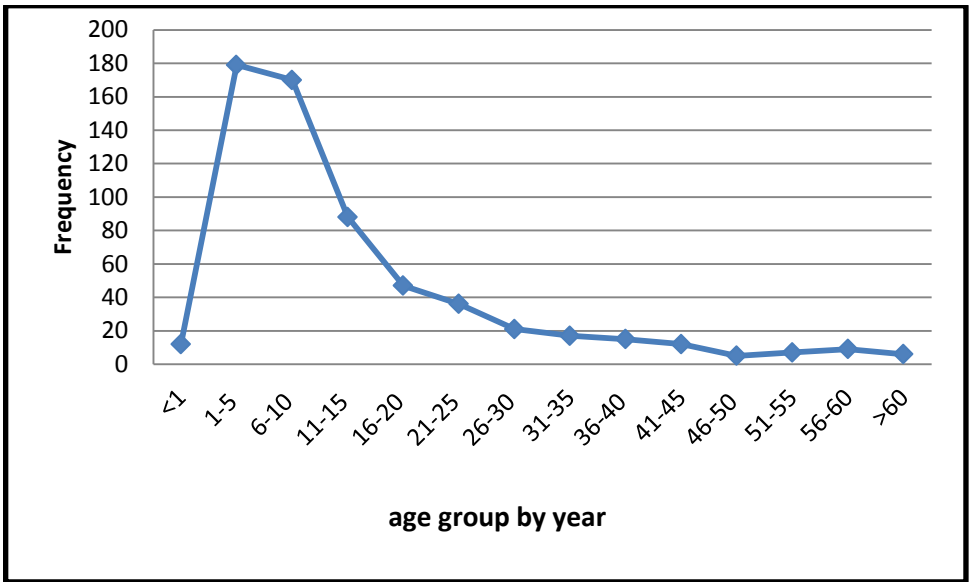


Figure 4-3: The numbers of ALL patients distributed according to age (5 years Interval). Note the beak at ~1-10 years old.

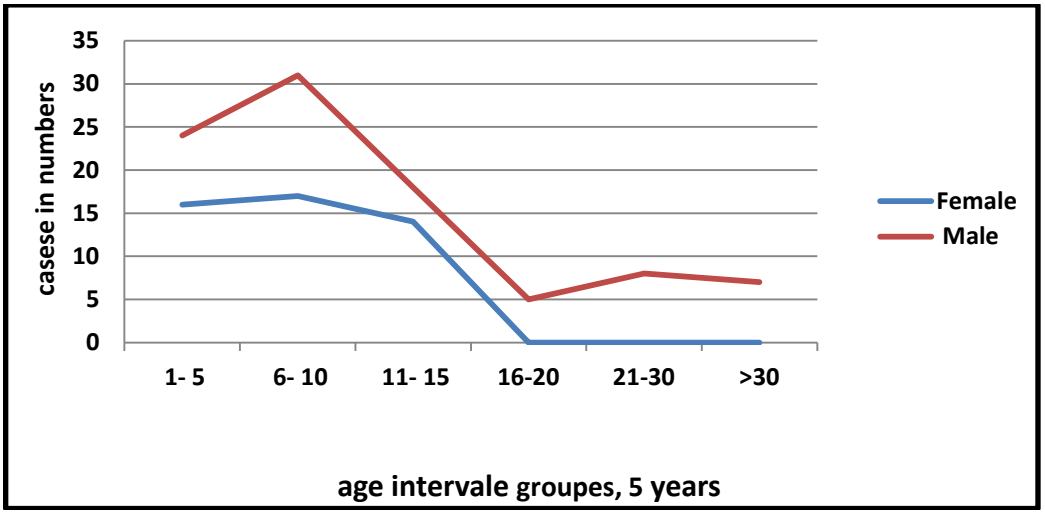


Figure 4-4: The number of ALL patients distributed according to gender. Note the female gender was not reported in more than 20 years' age group.

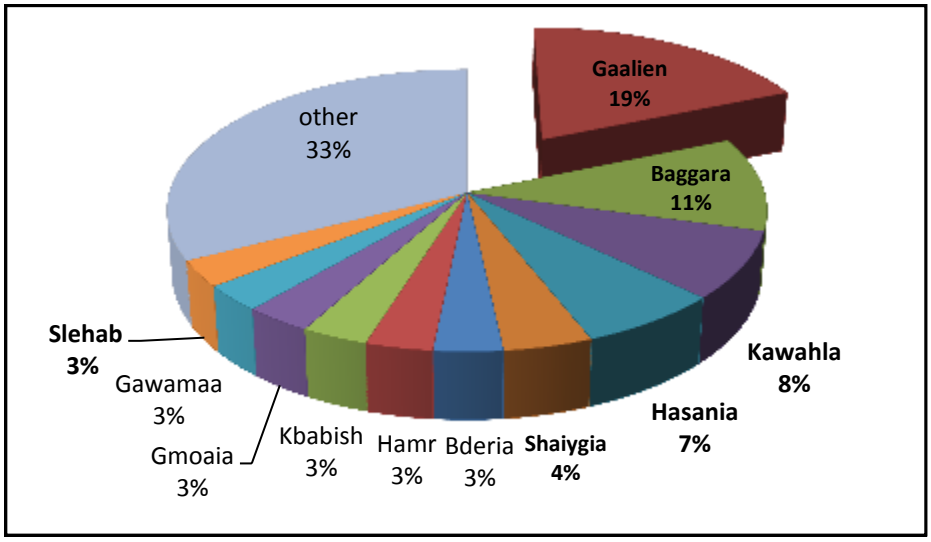


Figure 4-5: Relative frequencies of ALL of Afro-Asiatic Sudanese ethnic groups.

Table 4-3: Linguistic Classification of Sudanese ethnic groups included in this study.

Afro-Asiatic	**Nilo-Saharan**	**Niger Kordofanian**
-Beja:(Hadendowa,Beni-Amer and Rashaida) -Baggara:(Rizeigat, Meseria, BeniHelba, Haba niya, Ta`isha, Kinanah, Beni-Hessane)  - Gaalien, Hasania, Kawahla, Rufa`a, Batahin, Rekabia, Ashraf, Shukria, Shekhb, Halaween, Arkeen, Lahawin,Awamra,Shorab, Shaigiya, Robatab, Ababda, Bederia, Mnasir, Kababish, Ahamda, Mganeen, Hawawir, Hamr,Dar-Hamed, Bisharin, Benimora Shnabla, Hawazma, Goasma, Zeiyadiya, Gmoaia, Gawamaa, Geema, Gaafra and Beni-grar.	Nubian:(Mahas, Dnagla, Halfaw-ieenand Nuba) Fur, Tama, DajuFunj, Gumuz, Gemar Berta, Berti, Tunjur, Zaghawa and Masalit.	Fulani: (Falata,Hausa and Brno) Goraan, Borgo (Selehab) and Nuba Mountains.

Table 4-4: Hematological parameters between ALL Sudanese patients and healthy control group and correlation with GSTM1 and T1 genes.

Variable	WBC $x10^3/\mu l$	RBC $x10^6/\mu l$	HBg/dl	HCT%	MCV/fl	MCH /pg	LYM %	PLTx $10^9/L$	Blast%
Case: Mean±Sd	19±51	4±1	10±1	34±1	84±12	28±2	54±17	269±125	8±21
Control: Mean±Sd	8±2	5±5	12±1	36±3	81±3	27±1	39±9	353±8	00±0
P .value	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00
GSTM1: Mean±Sd									
Null	20 ±4	3.7 ±0.6	10.4±1	34±9	85±11	28.2	54±18	275±18	9±22
Present	19±6	3.7± 0.6	10.3±2	35±12	84±14	29±3	54±16	264±13	7±19
P.Value	0.84	0.62	0.82	0.64	0.59	0.07	0.60	0.28	0.42
GSTT1: Mean ±Sd									
Null	8±14	4±1	10.42	35±12	84±14	29±3	54±16	278±123	4±16
Present	26±62	3.7.±1	10.3±2	34±11	84±12	28±2	54±18	265±127	10±23
P.Value	0.04	0.96	0.62	0.66	0.83	0.19	0.52	0.96	0.81

Independent *t* –test and mean±SD.

.

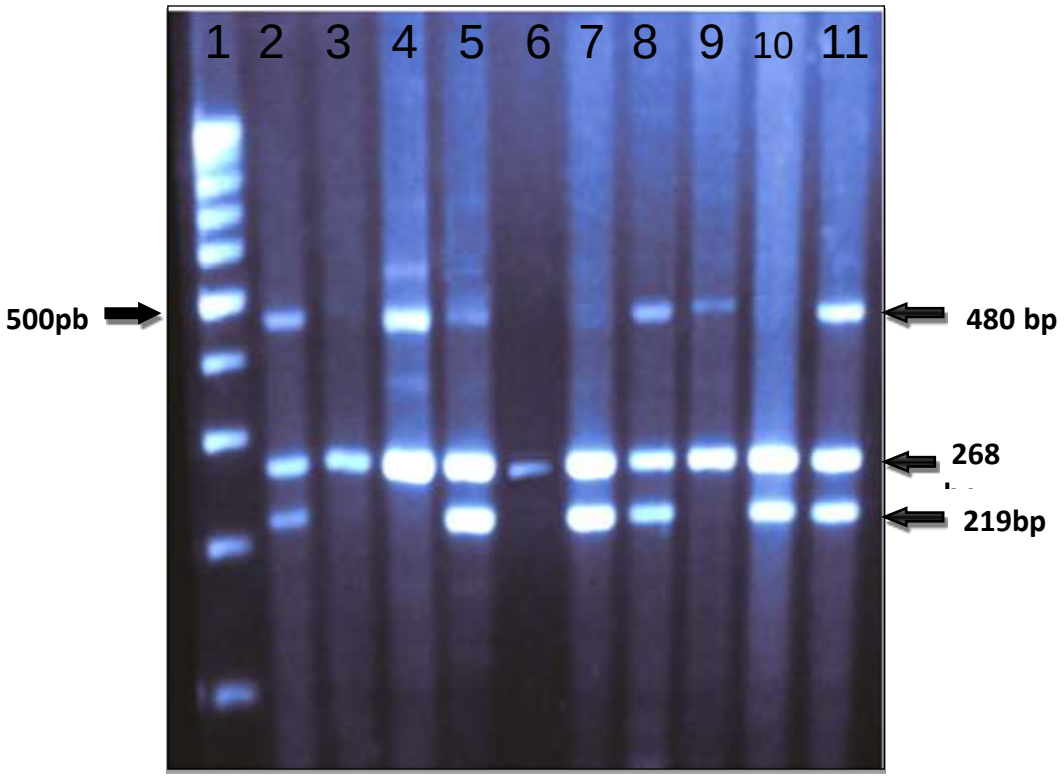


Figure 4-6: Analysis of GSTM1, T1 and P1 polymorphisms by multiplex PCR. Lane1: DNA molecular weight standard; Lanes 2, 5 and 11: *GSTT1* (480 bp) and *GSTM1* (219bp) homozygous present; Lanes 3 and 6 GSTT1 and GSTM1 double Null genotypes (homo zygous deletion, only the β-globin product (268 bp) is detected). Lane 7 and10 shows an individual with *GST M1* but *GSTT1* is Null. In lane 4and 9, *GSTM1* is Null, while a *GSTT1*is present

Table 4-5: *GSTM1*, *GSTP1* (rs1695-Exon5), *GSTT1* genotypes among acute lymphocytic Leukaemia patients and controls.

Genotype	Cases [*n*(%)]	Controls [*n*(%)]	Odd Ratio	95% CI Lower-Upper	*P.Value*
GSTP1(rs1695)					
AA	52(34.7)	86(61.1)	Rf		
AG	66 (44)	33 (25.4)	3.308	1.925-5.685	0.000
GG	32(21.3)	11(8.5)	4.811	2.23-10.3	0.000
AG+GG	98(65.5)	44(33.9)	3.684	2.24-6.043	0.000
GSTM1 :					
GSTM1-Present	72(48)	78(60)	1.72	1.1 -2.7	0.03
Null	78(52)	52(40)			
GSTT1 :					

GSTT1-Present	98(65)	91(70)	1.23	0.74 -2.0	0.44
Null	52(35)	39(30)			
Total	150(53.6)	130(46.4)			

AA: Normal homozygote IIe/IIe; AG; Heterozygote IIe/Val; GG: Mutant homozygote Val/Val .Chi-square

Table 4 -6: Association between *GSTT1*and *GSTM1*within the studied ALL patients.

	GSTM1		Total
	Null	**Present**	
GSTT1:			
Null count	21	31	52
%of total	14%	20.7%	34.7%
Present Count	57	41	98
%of Total	38%	27.3%	65.3%
Total Count	78	72	150
%of Total	52%	48%	100%

Table 4-7: Binary Regression for Risk Factors of ALL among the studied patients.

Variable	Significance Level	Exp	95% CI
			Lower-Upper
GSTT1	0.33	0.70	0.35-1.43
GSTM1	0.37	0.74	0.383-1.43
GSTP1	0.00	0.51	0.323-0.817
WBC	0.01	1.0	1.0- 1.10
Hb	0.00	2.4	1.86- 3.86

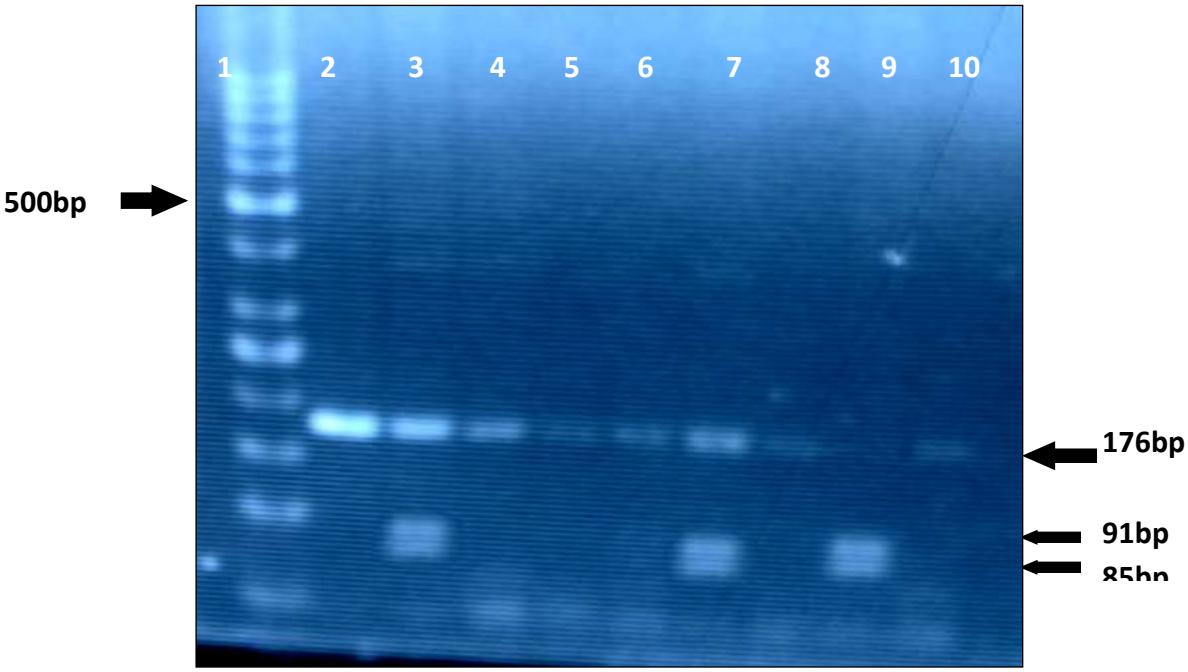


Figure 4-7: The PCR product of GSTP1 (176bp) digested with BsmAI enzyme cleaved to yield two fragments of 91 bp and 85bp sizes. Lane1: DNA molecular weight standard (100 bp), Lane: 2, 4, 5, 6, 8 &10: homozygous Wild type (ile105/ile 105); Lanes: 3 and 7 Heterozygous (ile105 / val 105). Lane: 9 homozygous mutant (val 105 /val 105)

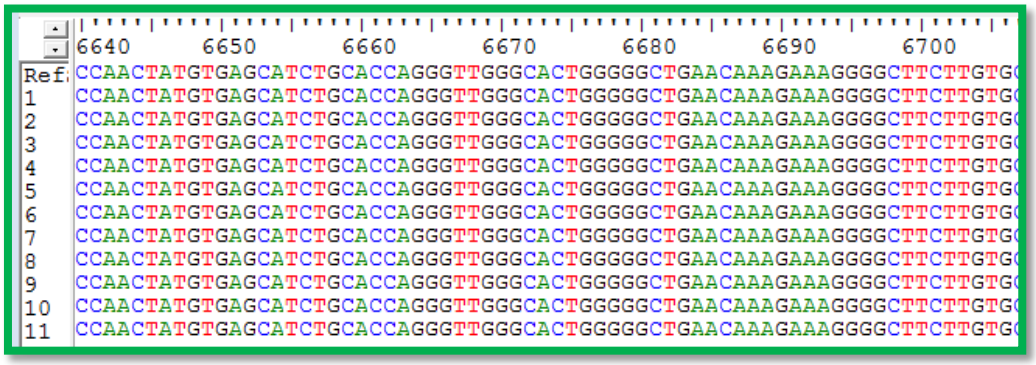


Figure 4-8: BioEdit multiple sequence alignment showed similarity of GSTP1 Sudanese cases with ALL when compared with reference sequence

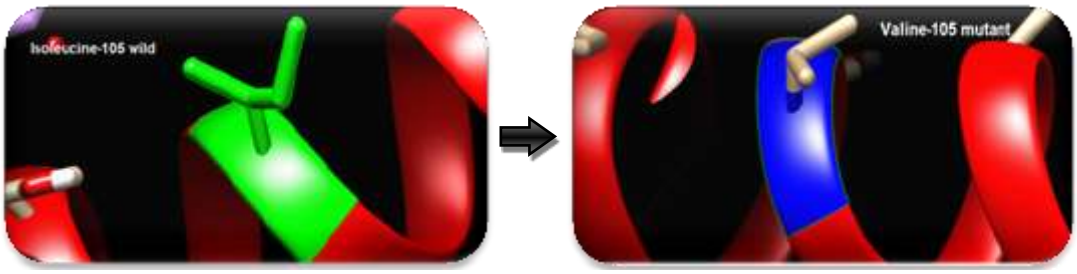


Figure 4-9: Bioinformatic analysis ofGSTP1 the isoleucine amino acid converted to valine at codon 105 using Chimera 1.8.

Table 4- 8: *NAT2* genotypes in healthy controls and Sudanese's ALL patients.

Genotype	Cases [n(%)]	Controls [n (%)]	Odd Ratio	95% CI Lower-Upper	*P.Value*
Tag (G590A) :					
GG	32(32)	51(46.4)	Rf		
AG	42(42)	42(38.1)	1.69	0.86 -2.94	0.13
AA	26(26)	17 (15.5)	2.43	1.14 -4.70	0.01
AA+AG	68(68)	59(53.6)	1.83	1.024- 3.22	0.03
Dde (A803G) :					
AA	31(31)	46(41.8)	Rf		
AG	37(37)	44(40)	1.24	0.664- 2.34	0.5
GG	32(32)	20(18.2)	2.374	1.15-4.88	0.02
AG+GG	69(69)	64(58.2)	1.60	0.906 –2.82	0.11
BamH1(G857A):					
GG	22(22)	41(37.3)	Rf		
AG	31(31)	36(32.7)	1.60	0.79- 3.20	0.2
AA	47(47)	33 (30)	2.65	1.34 -5.25	0.00
AG+AA	78(78)	69(62.7)	2.107	1.14- 3.88	0.01
Kpn(C481T):					
CC	25(25)	35(31.8)	Rf		
TC	45(45)	49(44.6)	1.28	0.66 -2.47	0.45
TT	30(30)	26(23.6)	1.61	0.77 -3.36	0.19
Total	100(52.4))	110(47.6)			

Rf (reference)

Table 4-9: Association of NAT2polymorphisms with acute lymphoblastic Leukaemia among Sudanese's patients

Variable	Significance Level	Exp	95% CI Lower-Upper
Dde A803G (rs1208)	0. 07	0.7 2	0. 501- 1.0 38

Tag G590A (rs1799930)	0.0 5	1.4 5	0. 995-2. 136
BamH1G857A(rs1801279)	0.02	1.48	1.048 -2. 106
Kpn1C481T (rs1799929)	0.28	1.23	0.839- 1.817

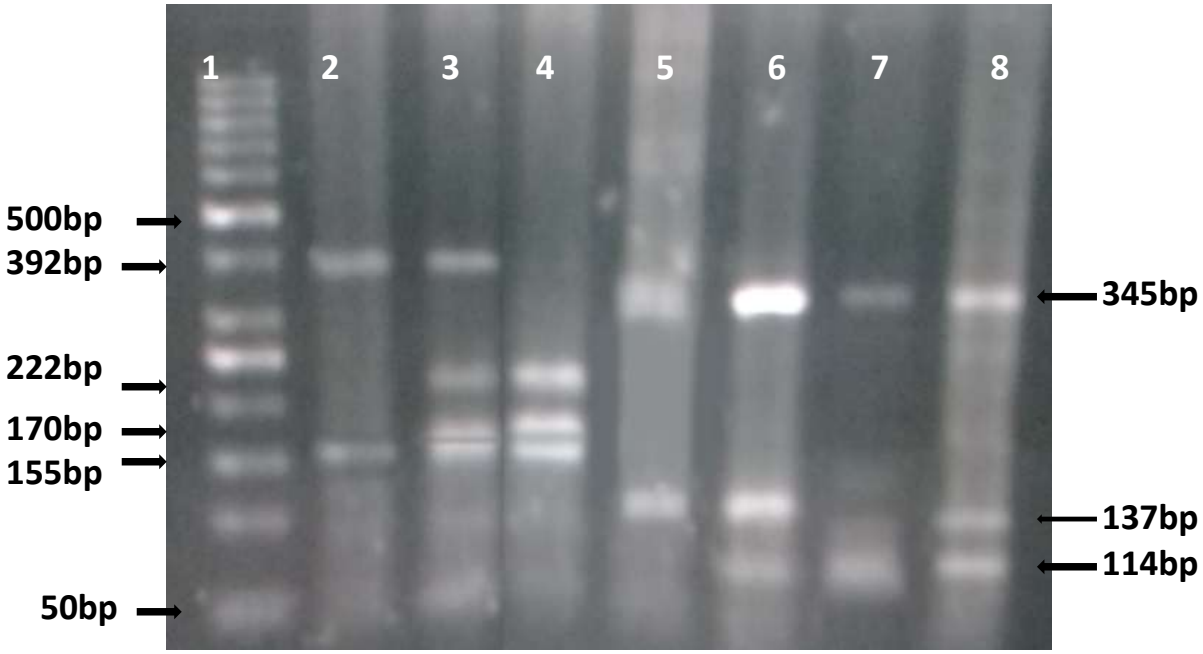


Figure 4-10: Tag G590A (rs1799930) digested PCR products of the amplified NAT2 region (547 bp) analyzed on 3% agarose. Lane1: DNA molecular weight standard (50bp), lane 2 is homozygote mutant (392 and155), lane 3 heterozygote (392,222,170 and 155bp) and lane 4 homozygote wild type (222, 170 and 155bp respectively). And Dde A803G (rs1208) digestion on lanes 5 is homozygote wild type (345 and 137bp), lane 6 and 8 are heterozygote (345,137 and 114bp) and lane 7 is homozygote mutant (345 and 114bp)

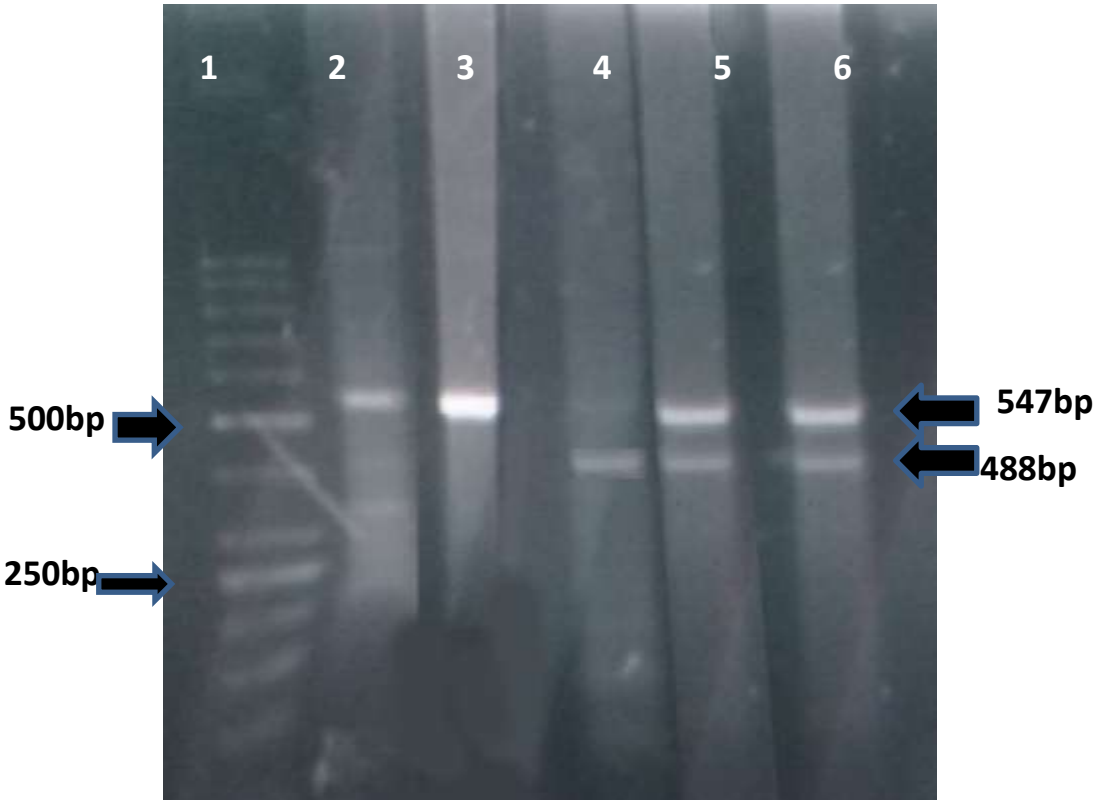


Figure 4-11: BamH1G857A (rs1801279) digestion of PCR product of the amplified NAT2 region analyzed on 3% agarose. Lane 1: DNA molecular weight standard (50bp), lane 2 and 3 are homozygote mutant (547bp), Lane 4 is homozygote wild type (488bp), and Lane5 and 6 are Heterozygote (547 and 488bp)

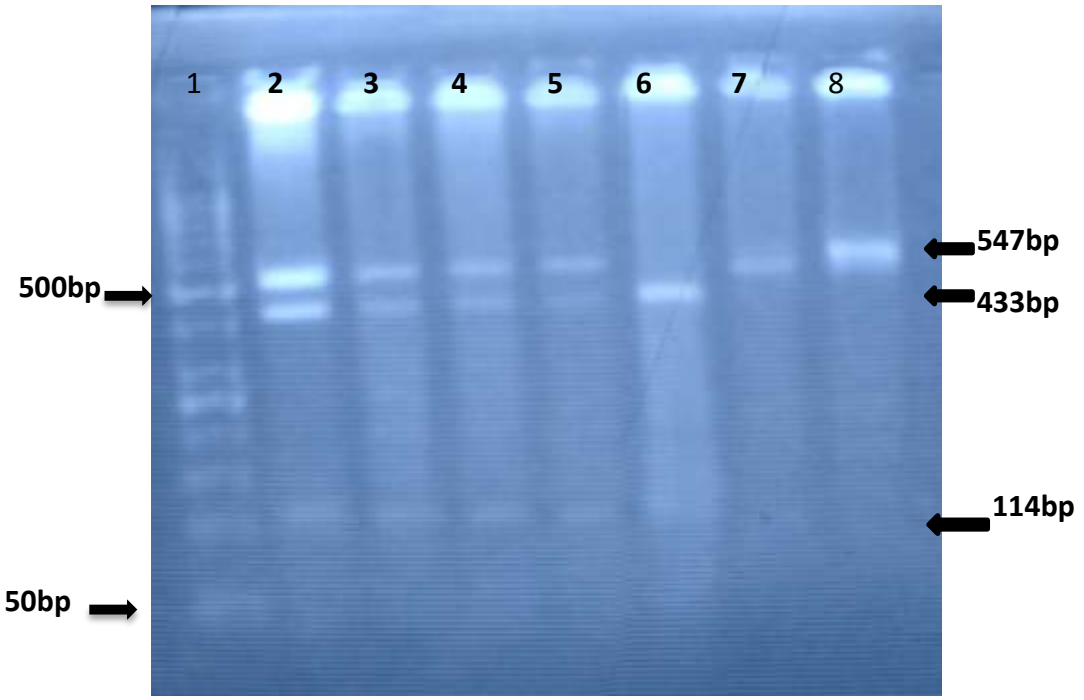


Figure 4-12:Kpn1C481T SNP (rs1799929) digestion of the NAT2 PCR product (547bp) analyzed of on 3% agarose. Lane 1: DNA molecular weight standard (50bp), lane 2, 3, 4 and 5 are heterozygote (547, 433and 114bp), Lane 6 is homozygote wild type (433and 114bp) and Lane7 and 8 are homozygote mutant (547bp).

.

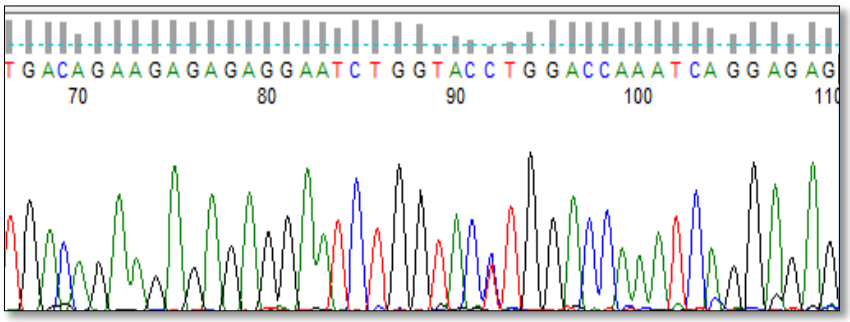


Figure 4-13: Sequences chromatogram viewed by Finch TV program, which showed heterozygosity (CT) in NAT2 at position 92.

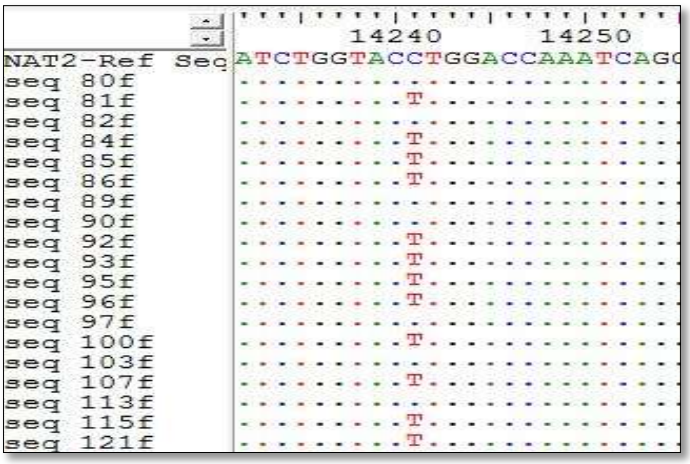


Figure 4-14: Multiple alignments for (NAT2547bp) Sudanese patient sequence forward (NO 80-100) and Sudanese control sequence(NO 103 and 121)with reference databases sequence using BioEdit Sequence ,show change of wild type (C) to mutant nucleotide (T) for Kpn1C481T (rs1799929) SNP at position 14241 in figure.

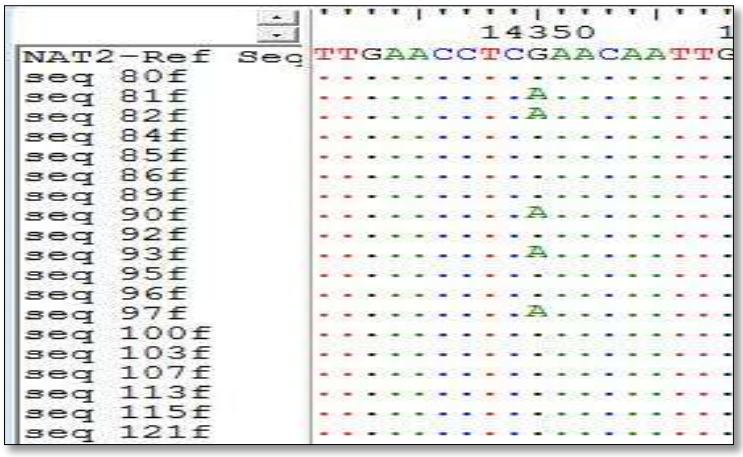


Figure 4-15: Multiple alignments for (NAT2547bp) patient sequence forward (NO 80-100) and control sequence (NO 103 -121) with reference databases sequence using BioEdit Sequence, show change of wild type (G) to mutant nucleotide (A) for Tag G590A (rs1799930)and BamH1 G857A (rs1801279) for cases sequences and no change to control sequences at position 143450 in figure.

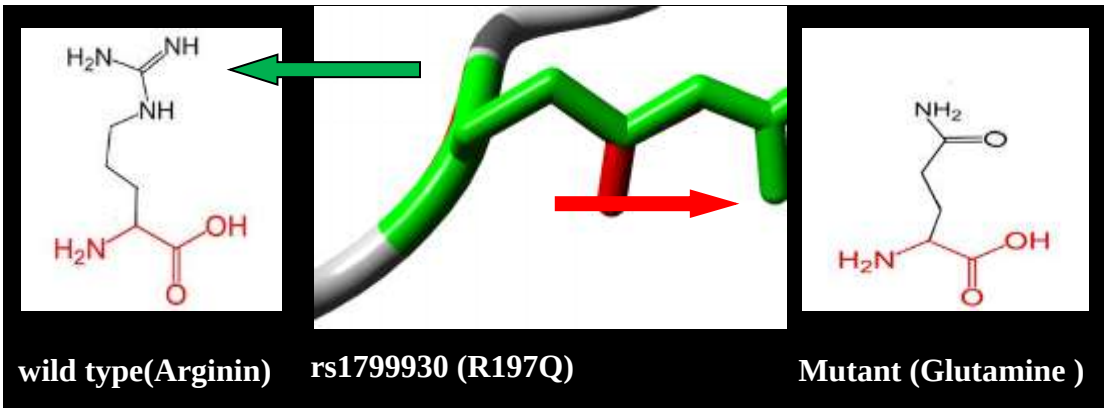


Figure 4-16. HOPE Modeling of the 3D structure of the NAT2 protein with selected SNPs. The protein (grey color) residues changed from wild type (green color) to mutant (red color).

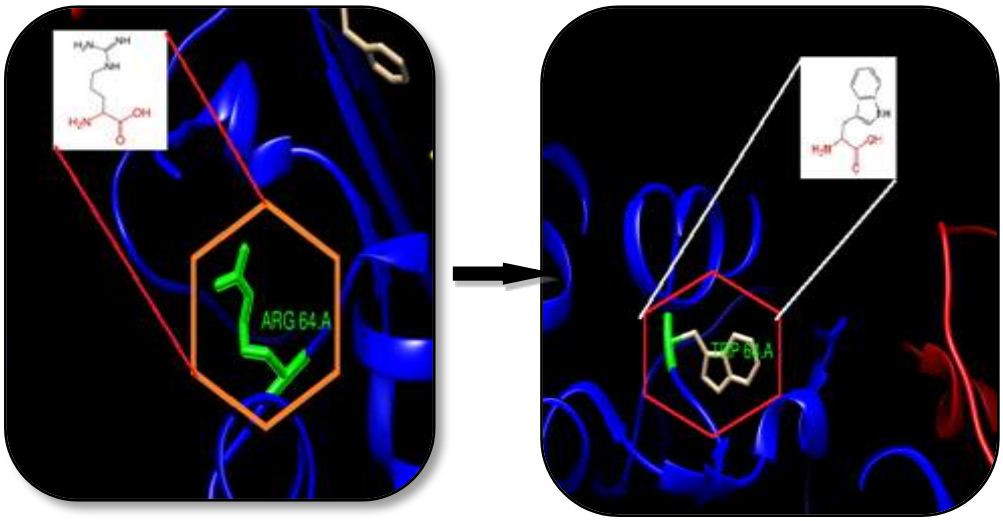


Figure 4-17: Chimera modelling of the native Arginine amino acid(R) substitution for a Tryptophan(W) at codon 64(rs1801279).

Chapter five

Discussion, Conclusions and Recommendations

Chapter five

5.1. Discusion:

Childhood Leukaemia is a widespread cancer in pediatric age that differs from adult cancers in its nature. The disease is one of the biggest challenges regarding worldwide mortality and morbidity, and shows considerable variation in disease incidence across geographical regions (Al-absi *et al*., 2017, Zheng *et al*., 2017).The incidence of malignancy has significantly risen recently in Sudan (Ahmed *et al*., 2014). This study showed that ALL is prevalent in younger age groups with a striking predominance of males when compared to females. In accordance with previous studies (Balta *et al*., 2003, Entezari *et al*., 2017, Saini *et al*., Errahhali *et al*., 2016), this difference (2:1female: male ratio) was the strongest among children age groups. There was a female occurrence among very young children until ~17 year and demonstrated zero rate above those years. This variation was strongest between children 6-10 years of age and younger. The feature rate peaks in the 1-10 year age group reproduced more than a 2-fold difference, which was not visible after 20years of age. And this result is partially in agreement with Ibrahim *et al*., 2017, in Sudanese patients with T-ALL (6-16 years: 58%).

Among the children groups (group two and three), an early peak between1-18year is observed un resembling age distribution of ALL in developing countries where a very marked early peak between 1-12 years (Saleh *et al*., 2017) has been observed. Interestingly, the majority of ALL cases occurred in individuals under the age of 20 years which is the most common malignancy in this age group (except in Africa and the Middle East) (Estey *et al*., 2014). The reported older ages herein are in concordance to another study in India where ALL age was most common at 6-14 years age old (Saini *et al*., Jain *et al*., 2017). Age-specific rates progressively decrease within the less than one-year age group (2.2%) infant. Infant age group is considered of poor prognosis, representing exclusive challenges to pediatric oncologists, because treatment can harmfully affect the children's growth, and can cause secondary malignancies. The risk then declines slowly until the mid-20s, and begins to rise again slowly after age 50 years. History of cancer was observed in about ~1% from cases in first relationship and about more than 60% of children the intake of specific types of nutrition (Soda water, falafel markets and sheep's) was documented. Markers expressed on cell-surface and cytoplasm grouped lymphoid lineage cell into B-cell

(80%) or T-cell (20%) subgroups that are suggestive to normal stages of lymphoid maturation (Estey *et al*., 2014). Parallel to that, in the current study B-cell was the most common phenotype (75%). Also the distribution of phenotypes among age groups is going with the theory denoting that T-ALL cell has the lowest percentages in ages lower than seven year and slightly increases in older ages. The morphologic recognition of lymphoblasts in the blood and bone marrow and their phenotypic characterization are of major importance to the correct diagnosis and classification of ALL. The current study denoted that these abnormal cells are arrested in all the new cases with significantly raised counts and lymphoblast (P =0.00 of both), however, some cases showed normal cells with features of leucopenia or thrompocytopenia, because they are under medication program. Also, variable degrees of anemia with low levels of HB and HCT were observed in all most cases, although the MCV and MHC appearance as normal, this may be due to folate drug intake by patients. In the current study.

Numerous molecular studies have attempted the establishment of significant information concerning the relationship of *GSTM1*, *GSTT1* and *GSP1* with cancer susceptibility. The mu and theta classes of *GST* isozymes have a common and broad range of substrate specificities, and they detoxify the reactive metabolites of exzenoboitic substances (Khabaz *et al*., 2016, Kaur, 2019). In the present study, the frequently highlighted genetic modifier polymorphisms of cancer risk *GSTM1*, *GSTT1* and *GSTP1* were screened among Sudanese ALL patients and matched controls. The study indicated that *GSTM1* is significantly more deleted among ALL Sudanese cases with slightly higher rate of individuals carrying double deletion for M1 and T1 in cases rather than control. The risk conferred to individuals who carry homozygous deletions in *GSTM1* appears to be high. These deletion variants are very helpful in epidemiological studies of cancer since they separate individuals into two distinct susceptibility groups: those who are able to detoxify possible carcinogens through metabolic pathways controlled by *GSTM1* and *GSTT1* genes and those who have lost that particular component of the detoxification strategies (Elhoseiny *et al*., 2014). In fact, the frequencies of homozygous *GSTM1* and *GSTT1* deletions carriers are extremely high (20%-50%) in the majority of studied populations to date (Uddin *et al*., 2014, Yang *et al*., 2013, Ada *et al*., 2012). *GSTM1* and possibly, *GSTT1* were thought to be concerned in the cancer at more than one site (Yang *et al*., 2013). Although, the risk conferred to individuals who express homozygous deletions in

GSTM1 or *GSTT1* seems to be very small, the degree of risk is increased when possible interactions of *GSTM1* and/or *GSTT1* with other factors are considered (Elhoseiny *et al*., 2014). While carriers of homozygous deletion in *GSTM1* and *GSTT1* genes are expected to lose enzyme activity (Ramzy *et al*., 2011) and thus predisposed to cancer, in a small number of investigations, it was also observed that *GSTM1* null genotypes showed to play a protective role for cancer (Yadav *et al*., 2010). The *GSTM1* null genotype rates varied from moderately to high rates in European and in Asians (38% to 67% and (33% to 63%, respectively) to lower rates in African and African-Americans populations (22% to 35%) (Klautau-Guimarães *et al*., 2005). On the other hand, *GSTP1*, the most existing isoform in the Leukaemia and is also concerned with response to oxidative stress (Senthilkumar and Thirumurugan, 2012) Ile105Val (rs1695) in P1 gene effects cytological toxicity and adapts the risk to work-related diseases. Apart from this, cancer and neuropathy, have been registered in people with this missense mutation (Basharat and Yasmin, 2017). Was reported to be higher in populations with African origin (42%) and lower in Asian and European origin populations (Balta *et al*., 2003, Zheng *et al*., 2017). In the current study, incidence of the mutant genotypes (ILe/Val and Val/Val) was significantly higher in ALL patients compared to controls, with fourfold increased risk of ALL. Most of these patients had an origin related to Khartoum and Darfur states; the states reporting the highest frequencies of ALL cases in Sudan.

Interestingly, this mutation affects cytological toxicity and adapts the risk to work-related diseases, cancer, neuropathy, NOx, SOx respiratory asthma, allergy etc. (Basharat and Yasmin, 2017). The fact that the polymorphic site at codon 105 is the active site for the catalytic property of GSTP1enzyme, might explain the wide range of possible expression alterations and sports its core role in xenobiotic Metabolism. It was also reported that heterozygote type Ile/Val was increased in a group of patients lower than 20 years when compared to patients at older age groups in CML patients (Elhoseiny *et al*., 2014).Applying the same assumption in our ALL patients indicated that the heterozygous type was lower in childhood group compared to adulthood patients. Authors suggested that the occurrence of valine allele confers elevated risk to ALL at later ages. The results of insilico analysis differ from wet lab, duo to the substitution of ‘Ile’ to ‘Val’ given a benign amino acid character, when using Sift software and then named as tolerated SNP, this variation depend on ethnic population. SNPs mutations of the *NAT2* gene had been descried to be an important risk factor of

various cancers (ALL, bladder, colon, head and neck, lung, breast…etc.), due to low or no *NAT2* activity. The results vary widely and are often discordant likely because of ethnic and geographic differences of the enrolled subjects (Buffler *et al*., 2005, Zheng *et al*., 2017, Hein and Doll, 2012, Di Pietro *et al*., 2012, Mittal *et al*., 2004).The consecutive study of the *NAT2* gene, has showed wide allelic diversity and inter-ethnic variation, which is reflected in the different types of acetylators present in the analyzed population (Isis Arias *et al*.,2014). A number of studies have tried to set up relations between polymorphic expression of GSTs classes and ALL risk in different ethnic populations (Zheng *et al*., 2017, Hunger and Mullighan, 2015, Chiaretti *et al*., 2014). Several wet lab studies highlighted the role of a number of the *NAT2* SNPs in carcinogenesis, some of them reported as slow acetylator phenotype leads to decrease the stability of protein Tag G590A (rs1799930), BamH1 G857A (rs1801279), Dde A803G (rs1208), the Kpn1 C481T (rs1799929) as silent and other SNPs diagnostic as fast acetylator phenotype (Santos *et al*., 2016, Zupa *et al*., 2009, Hein and Doll, 2012, Hein *et al*., 2000).

The main findings in our study, three of these SNPs (rs179 9930, rs1801279 and rs1208), were revealed to be significant in Sudanese patients with ALL (*P*=0.00), the strongest associations were with rs1799930 and rs1801279 and these genetic effects seem to be additive (*P*=0.05 and 0.02). While no significant differences were reported for the for the SNP (rs1799929) between case and controls in the different genotypes of Sudanese ALL patients, suggesting absence of possible effects on protein synthesis and supporting the reported data regarding that last SNP. Both homozygote and heterozygote genotypes of the Tag G590A and the BamH1 G857A SNPs were suggested to be associated with ALL etiology in. While only the homozygous state (GG) of the Dde A803G SNP reached the level of significance in the association analysis (P=0.02). The detected *NAT*2 inconsistency in the allelic frequency, can be related to carcinogens is via changes in the metabolic rate after contact to carcinogens (Adole *et al*., 2016). The allele frequencies of A (G590A) and G (A803G) have two times associated difference compared to one time in the case of A (G857A) with ALL Sudanese's patient. The molecular homology modeling bioinformatics methods, including SNP locations and docking application, have enhanced the knowledge of the *NAT*2 protein structure -function association (Abodlaa *et al*., 2019, Alabid *et al*., 2016). In this is study, a computational examination of the NAT2 SNPs using various publicly available softwares and databases was conducted to understand their possible

influence on the protein structure and function. Of these, two SNPs (rs1801279 (R64Q) and rs1799930 (R197Q)) were revealed to be deleterious, probably damaging and could decrease the stability of protein according to Sift, polyPhen-2 and the I-Mutant 3.0 softwares respectively. While the two other SNPs (rs1799929 and rs1208) were predicted as synonymous, tolerated and benign effects by three softwares (i.e. they do not alter the phenotype). These findings are generally in agreement with other studies(Arias *et al*., 2014, Hein *et al*., 2000). Despite in silico predictions, the rs1208 SNP was found to be significantly associated with ALL in wet lab analysis. The acetylator phenotype may be an outcome of other interacting genetic variations. Using ProjecHOPE software, the mutated residue of the R64Q SNP (rs1801279) was predicted to be smaller in size and neutral in charge when compare the positively charged wild-type. This variation in charge could disturb the ionic interaction bonds made by the native amino acid, while the size difference might lead to changes of the position of the new residue in the protein, preventing the creation of the same hydrogen bonds made by the original wild-type amino acid. Due to the difference in size, the mutation will cause the formation of an empty area in the core of the protein. The analysis revealed that the *NAT2* gene is highly polymorphic and can contribute to disease formation following changes in its amino acids. The results of this analysis agree with several studies(Arias *et al*., 2014, Shepard *et al*., 2011). About 20% of ALL children in developing countries either relapse or do not react to treatment which might be concerning to the occurrence of polymorphisms having an effect on the action of drug metabolizing enzymes and the response to medication (Basharat and Yasmin, 2017).

Sudan is an African country with great cultural diversity; it is a complex process to categorize the fixed cultures of the various people. These traditional societies have different social and cultural characteristics and distributed across wide geographical areas (DeSA, 2013). This variation is clearly evident in the number of tribes that had been incorporated in the study. The Afro-Asiatic group was recorded as the heights linguistic group (72%), with no correlation between this linguistic group and the NAT genotypes, largely due to the high number of tribes in that cluster. This may be due to the Arab tribes' presence, which is estimated at 70%, representing a major ethnic affiliation of the Sudanese population. Other linguistic group include the Arabized ethnic groups of Nubians, Zaghawa, and Copts (Greenberg and Kemmer, 1990,

DeSA, 2013). Within this cluster, the Ja`alin ethnic tribe becomes the most affected tribe, followed by Baggara (9%) (This is a group of tribes including Rizeigat, Meseria, Beni-Helba, Habaniya and Beni-Hessan). Kawahla and Hasania, living in the same area (White Nile) represented 8%. According to the geographical distribution, the central and Western Sudan states were associated with higher frequencies of ALL. This might be expected as Afro-Asiatic groups (Jaalin, Baggara, Kawahla and Hasania) were found to be the major inhibiting groups in those areas. Nilo-Saharan group (the four tribes live in the East and, Mahas and Dnagla from North of Sudan) are also of moderately high rates of ALL. Environmental and/or genetic factors are likely to play important roles in such distribution patterns; however, this role is rendered unclear due to the difficulties hampering disease etiology evaluation, importantly, shortage of studies and weak registry system. Interestingly, Dobon *et al* study showed shared ancestry with Afro-Asiatic and Nilo-Saharan Sudanese populations highlighting possible common genetic factors (Dobon *et al*., 2015).

The high rate of ALL cases (28%) in Khartoum, the capital city, may be due to the location of oncology services and health care centers. Also the recent increase of the rate of migration to the capital might have contributed to the reported higher rates. Likewise, whatever is the percentage of AL-Jazira state; it was mostly not the actual number, because there is another oncology center there. Generally, incidence fluctuates according to the different states of Sudan, and the association between the distribution of ALL patients and the common ethnic groups in the states is significant. Environmental exposure to carcinogens, genetic factors, or the life style adopted are very important in disease etiology, but environmental factors involving viral infection are more reliable in this childhood malignancy, as the rates of global infections of these viruses reach almost 90% of the general population (Loutfy *et al*., 2017, Mohamed *et al*., 2012). Several reports presented extensive data that highlight the relation between childhood malignancy mainly ALL, and viral infection, particularly herpes viruses (Cytomegalovirus and Epstein-Barr virus) (Alibek *et al*., 2013, Muhsin *et al*., 2014, Sehgal *et al*., 2010). ALL has not been eventually related to chemical contact, however, childhood ALL perhaps linked with maternal occupations related to exposure to pesticides during pregnancy. This might be the most rational explanation to the observed high frequencies of childhood category (1-6 year) occupying central state, also the center of major agricultural activities in Sudan (Alnil-Alabid, Shmal-

Kordofan and Algazera). Moreover, the high concordance between the geographic distribution of the virus in Sudan (Elawad *et al*., 2016), and ALL patients across Sudan states was observed. Kordofan and Darfur reported high rates after Khartoum state. At the level of ethnic groups also the Baggara ethnic group, living in these areas, was the second infected group after Ja`alin. This, in turn, indicates that Sudan might be experiencing a cancer outbreak, which carries many challenging characteri stics especially those related to environmental risk factors. Thirty percent of cancers in these countries are related to infection, and most cancer patients are young (Awadelkarim *et al*., 2012). Infectious agents could play a role in the development of childhood Leukaemia by the direct transforming ability of transforming viruses or by abnormal immunological responses to congenital, neonatal, or post-neonatal infections, which in turn promote secondary genetic or immunological alterations (Loutfy *et al*., 2017). This would not only indicate childhood cancer epidemic, but also would predict several infection associated cancers were on the rise.

5.2. Conclusion:

In conclusion, ALL has different features of mean age (6-10 years) and high incidence of Philadelphia chromosome (4.3%) in Sudanese children. While, in keeping with published literature, ALL shows a predominance of males, the B-phenotype frequency rate (%75.5) and the variation in haematological parameters. Our result documented clearly that *GSTM1* deletion and *GSTP1* genotype exhibits significant association with the risk of developing ALL.

Our findings considered that the age-related ALL incidences are affecting the adolescent patients compared with children and old age group, with a wide peak occurring at 1-10years. The higher rate in male than female.The genetic effect of polymorphism of *GSTM1*and *P1* and *NAT2* (Tag, Dde and BamH1) was reported in Sudanese's population with ALL. The Western and Central Sudan recorded the highest rates of ALL and most of ALL patients are from the Afro-Asiatic ethnic group.

5.3. Recommendations

- ✓ Increase the researches of more genetic and environmental factors which related to leukemia in Sudan.
- ✓ Use the huge data in database to determine the cancer etiology and development of more effective biologically-based therapies and guide national health screening strategies.

Reffrences

1.
2. Abdalla, A. H. S. (2016). Determination of Serum Zinc and Copper Levels in Sudanese Patients with Acute Lymphoblastic Leukemia, Khartoum State Hospitals, Sudan (2015). University of Gezira -thesis.
3. Abodlaa, A. E. E., Mursi, D., Khaier, M. A. M., Rahmanmasri, M. A., Husain, N. E. O. S. A. & Elnasri, H. A,(2019). Prediction of Structural and Functional Effects of Single Nucleotide Polymorphisms in NAT2gene, a Computational Analysis. International Journal of Genetics and Genomics, 7, 88.
4. Ada, A. O., Kunak, S. C., Hancer, F., Soydas, E., Alpar, S., Gulhan, M. & ISCAN, M. (2012). Association between GSTM1, GSTT1, and GSTP1 polymorphisms and lung cancer risk in a Turkish population. Molecular biology reports, 39, 5985-5993.
5. Adole, P. S., Kharbanda, P. S. & SHARMA, S, (2016). N-acetyltransferase 2 (NAT2) gene polymorphism as a predisposing factor for phenytoin intoxication in tuberculous meningitis or tuberculoma patients having seizures-A pilot study. The Indian journal of medical research, 143, 581.
6. Ahmad, O. B., Boschi-Pinto, C., Lopez, A. D., Murray, C. J., Lozano, R. & Inoue, M. (2001). Age standardization of rates: a new WHO standard. Geneva: World Health Organization, 9, 10.
7. Ahmed, H. G., Doumi, M. A., Eltybe, M. M., Mahmoud, T. A. & Ebnoof, S. O. (2014). Burden of cancer in north sudan: a community-based survey. Management in Health, 18.
8. Ahmed, H. G., Osman, S. I. & Ashankyty, I. M. (2012). Incidence of Epstein-Barr virus in pediatric leukemia in the Sudan. Clinical Lymphoma Myeloma and Leukemia, 12, 127-131.
9. Ahmed, H. G., Suliman, R. S. A. G., El aziz, M. S. A. & Alshammari, F. D. (2015). Molecular screening for Epstein Barr virus (EBV) among Sudanese patients with nasopharyngeal carcinoma (NPC). Infectious agents and cancer, 10, 6.
10. AL-Absi, B., Noor, S. M., Saif-Ali, R., Salem, S. D., Ahmed, R. H., Razif, M. F. & Muniandy, S. (2017). Association of ARID5B gene variants with acute

lymphoblastic leukemia in Yemeni children. Tumor Biology, 39, 1010428 317697573.

11. Alabid, T., Kordofani, A. A., Atalla, B., Altayb, H. N., Fadla, A. A., Mohamed, M., Osman, M. A. S. & Elamin, B. K. (2016). In silico Analysis of Single Nucleotide Polymorphisms (SNPs) in HumanVCAM-1 gene.
12. Alaqidi, A. A. H. & Alwash, M. M. (2019). Frequency Of GSTP1 (Ile105Val) Gene Polymorphism In Iraqi CML Patients And Its Association With Susceptibility To CML. Journal of Pharmaceutical Sciences and Research, 11, 819-823.
13. Alibek, K., Mussabekova, A., Kakpenova, A., Duisembekova, A., Baiken, Y., Aituov, B., Karatayeva, N. & Zhussupbekova, S. (2013). Childhood cancers: what is a possible role of infectious agents? Infectious agents and cancer, 8, 48.
14. Arias, I., Lecompte, N., Visbal, L., Curiel, I., Hernández, E., Garavito, P. & Silvera-Redondo, C. (2014). NAT2 gene polymorphisms in three indigenous groups in the Colombian Caribbean Coast region. Colombia Médica, 45, 148-153.
15. Awadelkarim, K., Arizzi, C., Elamin, E., Hamad, H., DE Blasio, P., Mekki, S., Osman, I., Biunno, I., Elwali, N. & Mariani- Costantini, R. (2008). Pathological, clinical and prognostic characteristics of breast cancer in Central Sudan versus Northern Italy: implications for breast cancer in Africa. Histopathology, 52, 445-456.
16. Awadelkarim, K. D., Mariani-Costantini, R. & Elwali, N. E. (2012). Cancer in the Sudan: an overview of the current status of knowledge on tumor patterns and risk factors. Science of the Total Environment, 423, 214-228.
17. Balta, G., Yuksek, N., Ozyurek, E., Ertem, U., Hicsonmez, G., Altay, C. & Gurgey, A. (2003). Characterization of MTHFR, GSTM1, GSTT1, GSTP1, and CYP1A1 genotypes in childhood acute leukemia. American journal of hematology, 73, 154-160.
18. Basharat, Z. & Yasmin, A. (2017). Energy landscape of a GSTP1 polymorph linked with cytological function decay in response to chemical stressors. Gene, 609, 19-27.
19. Bashasha, S., Kordofani, A., Osman, I., Musa, O. & Altayb, H. 2017. Prevalence of the different FAB sub type of Acute Myeloid Leukemia related to hematological parameters in Sudanese. J Hematol Blood Disord, 3, 102.

20. Birva, R., Vora, H., Trivedi, P. & Parikh, B.(2019). Flowcytometric Analysis of Leukemic Blasts-as Primary Screening Test for BCR/ABL1 Gene Rearrangement in B-ALL.
21. Black, R. E., Morris, S. S. & Bryce, J. (2003). Where and why are 10 million children dying every year? The lancet, 361, 2226-2234.
22. Blench, R. & Macdonald, K. (2006). The origins and development of African livestock: archaeology, genetics, linguistics and ethnography, Routledge.
23. Boattini, A., Castrì, L., Sarno, S., Useli, A., Cioffi, M., Sazzini, M., Garagnani, P., Defanti, S., Pettener, D. & Luiselli, D. (2013). mtDNA variation in East Africa unravels the history of afro- asiatic groups. American journal of physical anthropology, 150, 375-385.
24. Boothe, V. L., Boehmer, T. K., Wendel, A. M. & Yip, F. Y.(2014). Residential traffic exposure and childhood leukemia: a systematic review and meta-analysis. American journal of preventive medicine, 46, 413-422.
25. Borlak, J. & Reamon-buettner, S. M. (2006). N-acetyltransferase 2 (NAT2) gene polymorphisms in colon and lung cancer patients. BMC medical genetics, 7, 58.
26. Bouda, G. C., Traore, F., Couitchere, L., Raquin, M.-A., Guedenon, K. M., Pondy, A., Moreira, C., Rakotomahefa, M., Harif, M. & Patte, C. (2020). Advanced Burkitt Lymphoma in Sub-Saharan Africa Pediatric Units: Results of the Third Prospective Multicenter Study of the Groupe Franco-Africain d'Oncologie Pédiatrique. JCO Global Oncology, 5, 1-9.
27. Brisson, G. D., Alves, L. R. & Pombo-de-oliveira, M. S. (2015). Genetic susceptibility in childhood acute leukaemias: a systematic review. Ecancermedicalscience, 9.
28. Buffler, P. A., Kwan, M. L., Reynolds, P. & Urayama, K. Y. (2005). Environmental and genetic risk factors for childhood leukemia: appraising the evidence. Cancer investigation, 23, 60-75.
29. Chiaretti, S. & Foa, R. (2009). T-cell acute lymphoblastic leukemia. Haematologica.
30. Chiaretti, S., Zini, G. & Bassan, R. (2014). Diagnosis and subclassification of acute lymphoblastic leukemia. Mediterranean journal of hematology and infectious diseases, 6.

31. Chielle, E. O., Fortuna, P. C. & Maziero, J. S. (2016). Association between the glutathione S-transferase P1 (GSTP1) Ile105Val gene polymorphism in obese and overweight patients over 60 years. Jornal Brasileiro de Patologia e Medicina Laboratorial, 52, 211-216.
32. Churchman, M. L., Evans, K., Richmond, J., Robbins, A., Jones, L., Shapiro, I. M., Pachter, J. A., Weaver, D. T., Houghton, P. J. & Smith, M. A. (2016). Synergism of FAK and tyrosine kinase inhibition in Ph+ B-ALL. JCI insight, 1.
33. Coughlin, S. S. & Hall, I. J. (2002). Glutathione S-transferase polymorphisms and risk of ovarian cancer: a HuGE review. Genetics in Medicine, 4, 250.
34. Crawford, D. H. (2001). Biology and disease associations of Epstein–Barr virus. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 356, 461-473.
35. Crocco, P., Montesanto, A., Dato, S., Geracitano, S., Iannone, F., Passarino, G. & ROSE, G. (2019). Inter-individual variability in xenobiotic-metabolizing enzymes: implications for human aging and longevity. Genes, 10, 403.
36. Danker-Hopfe, H., Dasenbrock, C., Huss, A., Klaeboe, L., Mjönes, L., Moberg, L., Röösli, M., Scarfi, M., Van deventer, E. & Vanrongen, E. (2018). Recent Research on EMF and Health Risk, Twelfth report from SSM's Scientific Council on Electromagnetic Fields, 2017.
37. Desa, U. (2013). World population prospects: the 2012 revision. Population division of the department of economic and social affairs of the United Nations Secretariat, New York, 18.
38. Di pietro, G., Gadelha, S. R., Sousa, S. M. B., DE Melo, P. R. S. & SANTOS, F. R. (2012). N-acetyltransferase 2: Slow, intermediate or fast? A booming question of the molecular epidemiology in cancer research. Open Journal of Genetics, 2, 221.
39. Digiuseppe, J. A. (2016). Acute lymphoblastic leukemia/lymphoma: Diagnosis and minimal residual disease detection by flow cytometric immunophenotyping. Manual of molecular and clinical laboratory immunology, 207-216.
40. DimmendaaL, G. (2014). Marked nominative systems in Eastern Sudanic and their historical origin. Afrikanistik online, 11.
41. Dimmendaal, G. J. (2007). Eastern Sudanic and the Wadi Howar and VVadi el Milk Diaspora.

42. DimmendaaL, G. J. (2010). Differential object marking in Nilo-Saharan. Journal of African Languages and Linguistics, 31, 13-46.
43. Dimmendaal, G. J. 2011. Historical linguistics and the comparative study of African languages, John Benjamins Publishing.
44. Ding, F., LI, J.-P., Zhang, Y., QI, G.-H., Song, Z.-C. & YU, Y.-H. (2019). Comprehensive analysis of the association between the rs1138272 polymorphism of the GSTP1 gene and cancer susceptibility. Frontiers in physiology, 9, 1897.
45. Dobon, B., Hassan, H. Y., Laayouni, H., Luisi, P., Ricaño-Ponce, I., Zhernakova, A., Wijmenga, C., Tahir, H., Comas, D. & Netea, M. G. (2015). The genetics of East African populations: a Nilo-Saharan component in the African genetic landscape. Scientific reports, 5, 9996.
46. Dobritzsch, D., Grancharov, K., Hermsen, C., Krauss, G.-J. & Schaumlöffel, D. (2019). Inhibitory effect of metals on animal and plant glutathione transferases. Journal of Trace Elements in Medicine and Biology.
47. Dunna, N. R., Vuree, S., Kagita, S., Surekha, D., Digumarti, R., Rajappa, S. & Satti, V. (2014). Association of GSTP1 gene (I105V) polymorphism with acute leukaemia. Journal of genetics, 93, 60-63.
48. Dzian, A., Halasova, E., Matakova, T., Kavcova, E., Smolar, M., DOBROTA, D., Hamzik, J. & Mistuna, D. (2012). Lung adenocarcinoma and squamous cell carcinoma in association with genetic polymorphisms of GSTs in Slovak population. Neoplasma, 59, 160-167.
49. Eiman, A. (2018). Elshaygi."Assessment of Serum Levels of Electrolytes and Trace Elements in Leukaemia Patients in Sudan". Acta Scientific Pharmaceutical Sciences, 2, 136-141.
50. Eissa, D. S. & Ahmed, T. M. (2013). C677T and A1298C polymorphisms of the methylenetetrahydrofolate reductase gene: effect on methotrexate-related toxicity in adult acute lymphoblastic leukaemia. Blood Coagul Fibrinolysis, 24, 181-8.
51. Elawad, H., Kafi, S. & Elrahman, I (2016). The Possible Involvement of Epstein-Barr Virus in The Etiology of Leukemia. Cancer Prog Diagn, 1, 5-8.
52. Elhoseiny, S., El-Wakil, M., Fawzy, M. & Rahman, A. A. (2014). GSTP1 (Ile105Val) gene polymorphism: risk and treatment response in chronic myeloid leukemia. Journal of Cancer Therapy, 5, 1-10.

53. Entezari, M., Khatamsaz, S. & Dehghani, H (2017). Anti-proliferative effects curcumin in human acute lymphoblastic leukemia cell line. Journal of Paramedical Sciences, 8, 1-5.
54. Erdmann, F., Feychting, M., Mogensen, H., Schmiegelow, K. & Zeeb, H. (2019). Social Inequalities along the Childhood Cancer Continuum: an Overview of Evidence and a Conceptual Framework to Identify Underlying Mechanisms and Pathways. Frontiers in public health, 7.
55. Errahhali, M. E., Errahhali, M. E., Boulouiz, R., Ouarzane, M. & Bellaoui, M. (2016). Distribution and features of hematological malignancies in Eastern Morocco: a retrospective multicenter study over 5 years. BMC cancer, 16, 159.
56. Estey, E. H., Faderl, S. & Kantarjian, H. M. (2014). Hematologic Maligna ncies: Acute Leukemias, Springer.
57. Fadhlaoui- Zid, K., Rodríguez- Botigue, L., NAOUI, N., Benammar-Elgaaied, A., Calafell, F. & Comas, D. (2011). Mitochondrial DNA structure in North Africa reveals a genetic discontinuity in the Nile Valley. American journal of physical anthropology, 145, 107-117.
58. Farasani, A. (2019). Genetic variants of glutathione S-transferase and the risk of acute myeloid leukemia in a Saudi population. Saudi journal of biological sciences, 26, 1525-1530.
59. Farivar, S. & Ghazimoradi, M. H. (2019). DNA and RNA extraction from low amount of blood volume. Forensic science international, 303, 109931.
60. Feki-Tounsi, M., Khlifi, R., Louati, I., Fourati, M., Mhiri, M.-N., Hamza-Chaffai, A. & Rebai, A. (2017). Polymorphisms in XRCC1, ERCC2, and ERCC3 DNA repair genes, CYP1A1 xenobiotic metabolism gene, and tobacco are associated with bladder cancer susceptibility in Tunisian population. Environmental Science and Pollution Research, 24, 22476-22484.
61. Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D., Piñeros, M., Znaor, A. & Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. International journal of cancer, 144, 1941-1953.
62. Ferrari, A., Brecht, I. B., Gatta, G., Schneider, D. T., Orbach, D., Cecchetto, G., Godzinski, J., Reguerre, Y., Bien, E. & Stachowicz-Stencel, T. (2019). Defining and listing very rare cancers of paediatric age: consensus of the Joint

Action on Rare Cancers in cooperation with the European Cooperative Study Group for Pediatric Rare Tumors. European Journal of Cancer, 110, 120-126.

63. Gasic, V., Zukic, B., Stankovic, B., Janic, D., Dokmanovic, L., Lazic, J., Krstovski, N., Dolzan, V., Jazbec, J. & Pavlovic, S.(2018). Pharmacogenomic markers of glucocorticoid response in the initial phase of remission induction therapy in childhood acute lymphoblastic leukemia. Radiology and oncology, 52, 296-306.
64. Gatedee, J., Pakakassama, S., Muangman, S. & Pongstaporn, W. (2007). Glutathione S-transferase P1 genotypes, genetic susceptibility and outcome of therapy in thai childhood acute lymphoblastic leukemia. Asian Pacific Journal of Cancer Prevention, 8, 294.
65. Gatt, M. E. & Izraeli, S. (2019). Lymphoid leukemias. Clinical Immunology. Elsevier.
66. Genes, C. & Isruption, D. (2018). Genetic Epidemiology of Cancer. Textbook of Cancer Epidemiology, 81.
67. Gill, R. D., Martin, C. J. & Sanchez-Felix, M. V. (2016). Computer-assisted means for assessing lifestyle risk factors. Google Patents.
68. Gómez-Almaguer, D., Marcos-Ramírez, E. R., Montaño-Figueroa, E. H., Ruiz-Arguelles, G. J., Best-Aguilera, C. R., Del Carmen López-Sánchez, M., Barrera-Chairez, E., López-Arrollo, J. L., Ramos-Peñafiel, C. O. & León-Peña, A. (2017). Acute leukemia characteristics are different around the world: the Mexican perspective. Clinical Lymphoma Myeloma and Leukemia, 17, 46-51.
69. Gondos, A., Brenner, H., Wabinga, H. & Parkin, D.(2005). Cancer survival in Kampala, Uganda. British journal of cancer, 92, 1808.
70. Goodrich, J. M. & Basu, N. (2012). Variants of glutathione s-transferase pi 1 exhibit differential enzymatic activity and inhibition by heavy metals. Toxicology in Vitro, 26, 630-635.
71. Gra, O., Mityaeva, O., Berdichevets, I., Kozhekbaeva, Z., Fesenko, D., Kurbatova, O., Goldenkova-Pavlova, I. & Nasedkina, T. (2010). Microarray-Based Detection of CYP1A1, CYP2C9, CYP2C19, CYP2D6, GSTT1, GSTM1, MTHFR, MTRR, NQO1, NAT2, HLA-DQA1, and AB0 Allele Frequencies in Native Russians. Genetic testing and molecular biomarkers, 14, 329-342.
72. Gra, O. A., Glotov, A. S., Nikitin, E. A., Glotov, O. S., Kuznetsova, V. E., Chudinov, A. V., Sudarikov, A. B. & Nasedkina, T. V. (2008). Polymorphisms

in xenobiotic- metabolizing genes and the risk of chronic lymphocytic leukemia and non- Hodgkin's lymphoma in adult Russian patients. American journal of hematology, 83, 279-287.

73. Greenberg, J. H. & Kemmer, S. (1990). On language: selected writings of Joseph H. Greenberg, Stanford University Press.
74. Guaoua, S., Ratbi, I., Laarabi, F. Z., Elalaoui, S. C., Jaouad, I. C., Barkat, A. & Sefiani, A. (2014). Distribution of allelic and genotypic frequencies of NAT2 and CYP2E1 variants in Moroccan population. BMC genetics, 15, 156.
75. Hamad, H. (2006). Cancer initiatives in Sudan. Annals of oncology, 17, viii32-viii36.
76. Hayes, R. B., Songnian, Y., Dosemeci, M. & Linet, M. (2001). Benzene and lymphohematopoietic malignancies in humans. American journal of industrial medicine, 40, 117-126.
77. Hein, D. W. & Doll, M. A. (2012). Accuracy of various human NAT2 SNP genotyping panels to infer rapid, intermediate and slow acetylator phenotypes. Pharmacogenomics, 13, 31-41.
78. Hein, D. W., Doll, M. A., Fretland, A. J., Leff, M. A., Webb, S. J., Xiao, G. H., Devanaboyina, U.-S., Nangju, N. A. & Feng, Y. (2000). Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms. Cancer Epidemiology and Prevention Biomarkers, 9, 29-42.
79. Henn, B. M., Botigue, L. R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J. K., Fadhlaoui-Zid, K., Zalloua, P. A., Moreno-estrada, A. & Bertranpetit, J. (2012). Genomic ancestry of North Africans supports back-to-Africa migrations. PLoS genetics, 8, e1002397.
80. Hunger, S. P. & Mullighan, C. G. (2015). Acute lymphoblastic leukemia in children. New England Journal of Medicine, 373, 1541-1552.
81. Hussen, M. M. A. (2019). Association of Cytochrome P450 2 E1 (C1053T) and NADPH Quinone Oxide Reducatase 1 (C609T)(C 465T) Genes Polymorphism with Acute Lymphoblastic Leukemia in Sudanese Patients. Sudan University of Science & Technology.
82. Ibrahim, A. S., Khaled, H. M., Mikhail, N. N., Baraka, H. & Kamel, H. (2014). Cancer incidence in Egypt: results of the national population-based cancer registry program. Journal of cancer epidemiology, 2014.

83. Ibrahim, M. E. & Osman, E.-A. I. (2011). Myeloid Leukemia: A Molecular Focus on Etiology and Risk Within Africa. Myeloid Leukemia-Basic Mechanisms of Leukemogenesis. InTech.
84. Ikpatt, O., Kuopio, T., Ndoma-Egba, R. & Collan, Y. (2002). Breast cancer in Nigeria and Finland: epidemiological, clinical and histological comparison. Anticancer research, 22, 3005-3012.
85. Inskip, P. D., Veiga, L. H., Brenner, A. V., Sigurdson, A. J., Ostroumova, E., Chow, E. J., Stovall, M., Smith, S. A., Leisenring, W. & Robison, L. L. (2019). Hyperthyroidism After Radiation Therapy for Childhood Cancer: A Report from the Childhood Cancer Survivor Study. International Journal of Radiation Oncology* Biology* Physics, 104, 415-424.
86. Jain, S., Jain, S., Kapoor, G., Virmani, A. & Bajpai, R. (2017). No impact of disease and its treatment on bone mineral density in survivors of childhood acute lymphoblastic leukemia. Pediatric blood & cancer, 64, e26271.
87. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E. & Forman, D. (2011) . Global cancer statistics. CA: a cancer journal for clinicians, 61, 69-90.
88. Jemal, A., Bray, F., Forman, D., O'brien, M., Ferlay, J., Center, M. & Parkin, D. M. (2012). Cancer burden in Africa and opportunities for prevention. Cancer, 118, 4372-4384.
89. Jones, R. P. (2015). Roles for Cytomegalovirus in Infection, Inflammation, and Autoimmunity. Infection and Autoimmunity. Elsevier.
90. Joseph, T., Kusumakumary, P., Chacko, P., Abraham, A. & Radhakrishna Pillai, M. (2004). Genetic polymorphism of CYP1A1, CYP2D6, GSTM1 and GSTT1 and susceptibility to acute lymphoblastic leukaemia in Indian children. Pediatric blood & cancer, 43, 560-567.
91. Josephy, P. D. (2010). Genetic variations in human glutathione transferase enzymes: significance for pharmacology and toxicology. Human genomics and proteomics: HGP, 2010.
92. Kaur, G. (2019). Influence of GST polymorphism on treatment response in breast cancer patients.
93. Khabaz, M. N., AL-Maghrabi, J. A., Nedjadi, T. & Gar, M. (2016). Does Val/Val genotype of GSTP1 enzyme affects susceptibility to colorectal cancer in Saudi Arabia. Neuroendocrinol. Lett, 37, 46-52.

94. Khan, M. I., Naseem, L., Manzoor, R. & Yasmeen, N. (2017). Mortality Analysis in Children during Induction Therapy for Acute Lymphoblastic Leukemia. Journal of Islamabad Medical & Dental College, 6, 69-72.
95. Khlifi, R., Messaoud, O., Rebai, A. & Hamza-Chaffai, A. (2013). Polymorphisms in the human cytochrome P450 and arylamine N-acetyltransferase: susceptibility to head and neck cancers. BioMed research international, 2013.
96. Klautau-Guimarães, M. D. N., Hiragi, C. D. O., D'ascencão, R. F., Oliveira, S. F., Grisolia, C. K., Hatagima, A. & Ferrari, Í. (2005). Distribution of glutathione S-transferase GSTM1 and GSTT1 null phenotypes in Brazilian Amerindians. Genetics and molecular biology, 28, 32-35.
97. Kussick, S. J. (2018). Flow Cytometric Principles in Hematopathology. Hematopathology. Elsevier.
98. Lewis, M. P. (2009). Ethnologue: Languages of the world, SIL international.
99. Lichtman, M. A., Beutler, E., Kipps, T. J., Seligsohn, U., Kaushansky, K. & Prchal, J. T. (2006). Williams hematology, McGraw-Hill New York.
100. Loutfy, S. A., Abo-ShadI, M. A., Fawzy, M., El-WakiL, M., Metwally, S. A., Moneer, M. M., Fattah, N. F. A., Kassem, S. & Elgebaly, A. (2017). Epstein-Barr virus and cytomegalovirus infections and their clinical relevance in Egyptian leukemic pediatric patients. Virology journal, 14, 46.
101. Lukas, C., Selinski, S., Prager, H.-M., Blaszkewicz, M., Hengstler, J. G. & GOLKA, K. (2017). Occupational bladder cancer: Polymorphisms of xenobiotic metabolizing enzymes, exposures, and prognosis. Journal of Toxicology and Environmental Health, Part A, 80, 439-452.
102. Mathers, C. D. & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. PLoS medicine, 3, e442.
103. Mathoulin-pelissier, S. & Pritchard-jones, K. (2019). Evidence-based data and rare cancers: The need for a new methodological approach in research and investigation. European Journal of Surgical Oncology, 45, 22-30.
104. Maude, S. L. & Hunger, S. P. (2018). Update in Pediatric Oncology: Section A-New Developments in the Treatment of Pediatric Acute Lymphoblastic Leukemia. Update in Pediatrics. Springer.

105. Maxwell, R. R. & Cole, P. D. (2017). Pharmacogenetic predictors of treatment-related toxicity among children with acute lymphoblastic leukemia. Current hematologic malignancy reports, 12, 176-186.
106. Mehta, P. (2011). Postgraduate Haematology. JAMA, 305, 1713-1714.
107. Miller, C. (2016). Language and ethnic statistics in 20th century Sudanese censuses and surveys.
108. Miller, D., Bryant, J., Madsen, E. & Ghiorse, W. (1999). Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Applied and environmental microbiology, 65, 4715-4724.
109. Mittal, R. D., Srivastava, D. S. & Mandhani, A. (2004). NAT2 gene polymorphism in bladder cancer: a study from North India. International braz j urol, 30, 279-285.
110. Mohamed, O. W., Bakhiet, S. M., Gerges, H. K. & Kordofani, A. A. (2017). Serum zinc levels in sudanese patients with acute lymphocytic leukemia. Int J Clin and Biomed Res, 3, 18-20.
111. Mohamed, W. S., Loutfy, S. A., Rashed, R. A., Elazhary, N. M. & Amer, M. M. (2012). Prospective study for CMV detection in cancer patients: comparison between PCR, antigenemia, and serological assays. Life Sci J, 9, 4536-41.
112. Muhsin, M. A., Jebur, M. S. & Hussan, B. M. (2014). Detection of Epsteinâ€“Barr virus and Cytomegalovirus in Iraqi Acute Lymphoblastic Leukemia patients. AL-qadisiyah medical journal, 10, 201-208.
113. Nadkarni, A. & Hariharan, P. (2019). Molecular Techniques for Prenatal. Hematopathology: Advances in Understanding, 469.
114. Nasr, A. S., Sami, R. M., Ibrahim, N. Y. & Darwish, D. O. (2015). Glutathione S transferase (GSTP 1, GSTM 1, and GSTT 1) gene polymorphisms in Egyptian patients with acute myeloid leukemia. Indian journal of cancer, 52, 490.
115. Nori, E. M. & Ibrahim, I. K. (2018). Polymorphism in the gluthathione s-transferase theta and mu genes and susceptibility to lymphoid leukemia in sudanese patients. European Journal of Biomedical, 5, 38-41.
116. Ogata, M., Satou, T., Kawano, R., Yoshikawa, T., Ikewaki, J., Kohno, K., Ando, T., Miyazaki, Y., Ohtsuka, E. & Saburi, Y. (2011). High incidence of cytomegalovirus, human herpesvirus- 6, and Epstein–Barr virus reactivation in

patients receiving cytotoxic chemotherapy for Adult T cell leukemia. Journal of medical virology, 83, 702-709.

117. Ohki, K., Kiyokawa, N., Saito, Y., Hirabayashi, S., Nakabayashi, K., Ichikawa, H., Momozawa, Y., Okamura, K., Yoshimi, A. & Ogata-Kawata, H. (2019). Clinical and molecular characteristics of MEF2D fusion-positive B-cell precursor acute lymphoblastic leukemia in childhood, including a novel translocation resulting in MEF2D-HNRNPH1 gene fusion. haematologica, 104, 128-137.
118. Osman, T. A. (2017). Association of GSTM1Null Gene Polymorphism with Patients of Chronic Lymphocytic Leukemia in Khartoum. Sudan University of Science & Technology.
119. Osman, T. A., Satti, A. A., BøE, O. E., Yang, Y.-H., Ibrahim, S. O. & Suleiman, A. M. (2010). Pattern of malignant tumors registered at a referral oral and maxillofacial hospital in Sudan during 2006 and 2007. Journal of cancer research and therapeutics, 6, 473.
120. Paietta, E. (2018). Immunobiology of acute leukemia. Neoplastic diseases of the blood. Springer.
121. Parkin, D. M., Sitas, F., Chirenje, M., Stein, L., Abratt, R. & Wabinga, H. 2008. Part I: Cancer in Indigenous Africans—burden, distribution, and trends. The lancet oncology, 9, 683-692.
122. Patterns, D. (2010). Acute Lymphoblastic Leukemia: Epidemiology. Adult Acute Lymphocytic Leukemia: Biology and Treatment, 77.
123. Paul, L. M., Simons, G. F. & Fennig, C. D. (2009). Ethnologue: Languages of the world. Dallas, TX: SIL International. Available online at www. ethnologue. com/. Retrieved June, 19, 2011.
124. Pavlovic, S., Kotur, N., Stankovic, B., Zukic, B., Gasic, V. & Dokmanovic, L. (2019). Pharmacogenomic and Pharmacotranscriptomic Profiling of Childhood Acute Lymphoblastic Leukemia: Paving the Way to Personalized Treatment. Genes, 10, 191.
125. Peres, M. A., Macpherson, L. M., Weyant, R. J., Daly, B., Venturelli, R., Mathur, M. R., Listl, S., Celeste, R. K., Guarnizo-herreño, C. C. & Kearns, C. (2019). Oral diseases: a global public health challenge. The Lancet, 394, 249-260.
126. Provan, D. & Gribben, J. (2010). Molecular hematology, John Wiley & Sons.

127. Pui, C.-H. (2011). Acute lymphoblastic leukemia, Springer.
128. Rafieemehr, H., Calhor, F., Esfahani, H. & G, S. G. (2019). Risk of Acute Lymphoblastic Leukemia: Results of a Case-Control Study. Asian Pacific journal of cancer prevention: APJCP, 20, 2477.
129. Ramzy, M. M., Solliman, M., Abdel-hafiz, H. A. & Salah, R. (2011). Genetic polymorphism of GSTM1 and GSTP1 in lung cancer in Egypt. Int J Collab Res Int Med Public Health (IJCRIMPH), 3, 41-51.
130. Reehan, N. S. R. (2018). A comparative study of oxidative stress and Some physiological and biochemical variables in patients with myeloid and lymphoid leukemia in Baghdad city. University of Anbar.
131. Roberts, K. G. (2017). The biology of Philadelphia chromosome-like ALL. Best Practice & Research Clinical Haematology, 30, 212-221.
132. Rodgers, K. M., Udesky, J. O., Rudel, R. A. & Brody, J. G. (2018). Environmental chemicals and breast cancer: an updated review of epidemiological literature informed by biological mechanisms. Environmental research, 160, 152-182.
133. Saeed, I. E., Weng, H. Y., Mohamed, K. H. & Mohammed, S. I. (2014). Cancer incidence in Khartoum, Sudan: first results from the Cancer Registry, 2009–2010. Cancer medicine, 3, 1075-1084.
134. Saeed, M. E., Cao, J., Fadul, B., Kadioglu, O., Khalid, H. E., Yassin, Z., Mustafa, S. M., Saeed, E. & Efferth, T. (2016). A five-year survey of cancer prevalence in Sudan. Anticancer research, 36, 279-286.
135. Saini, m., Barman, m., Ghanghoriya, p., Agarwal, g. & Singhothiya, l. Pattern of Childhood Malignancy in Central india: A three years retrospective study. ratio, 82, 17.5.
136. Saleh, R., Musa, H. H. & Hamid, G. A. (2017). EPIDEMIOLOGICAL STUDY OF ACUTE LYMPHOBLASTIC LEUKEMIA IN YEMEN. European Journal of Biomedical, 4, 794-797.
137. Santos, E. C. L. D., Pinto, A. C., Klumb, E. M. & Macedo, J. M. B. (2016). Polymorphisms in NAT2 (N-acetyltransferase 2) gene in patients with systemic lupus erythematosus. Revista brasileira de reumatologia, 56, 521-529.
138. Santovito, A., Burgarello, C., Cervella, P. & Delpero, M. (2010). Polymorphisms of cytochrome P450 1A1, glutathione s-transferases M1 and T1

genes in Ouangolodougou (Northern Ivory Coast). Genetics and molecular biology, 33, 434-437.

139. Satoh, T. (2007). Genetic polymorphism in drug metabolism and toxicity: Linking animal research and risk assessment in man. AATEX, 14, 443-445.

140. Scheinfeldt, L. B., Soi, S. & Tishkoff, S. A. (2010). Working toward a synthesis of archaeological, linguistic, and genetic data for inferring African population history. Proceedings of the National Academy of Sciences, 107, 8931-8938.

141. Sehgal, S., Mujtaba, S., Gupta, D., Aggarwal, R. & Marwaha, R. (2010). High incidence of Epstein Barr virus infection in childhood acute lymphocytic lukemia: A preliminary study. Indian Journal of Pathology and Microbiology, 53, 63.

142. Senthilkumar, K. & Thirumurugan, R. (2012). GSTM1 and GSTT1 allele frequencies among various Indian and non-Indian ethnic groups. Asian Pacific Journal of Cancer Prevention, 13, 6263-6267.

143. Shah, S., Nawaz, W., Jalil, B. & Khan, H. A. (2019). Classification of Normal and Leukemic Blast Cells in B-ALL Cancer Using a Combination of Convolutional and Recurrent Neural Networks. ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging. Springer.

144. Sharkey, H. J. (2008). Arab identity and ideology in Sudan: The politics of language, ethnicity, and race. African Affairs, 107, 21-43.

145. Shepard, G., Lawson, H., Hawkins, G. & Owen, J. (2011). BsaXI/RFLP analysis of initial or selectively reamplified PCR product is unreliable in detecting the V617F mutation in JAK2. International journal of laboratory hematology, 33, 267-271.

146. Singh, S. A. & Ghosh, S. K. (2019). Metabolic phase I (CYPs) and phase II (GSTs) gene polymorphisms and their interaction with environmental factors in nasopharyngeal cancer from the ethnic population of Northeast India. Pathology & Oncology Research, 25, 33-44.

147. Sitas, F., Parkin, D. M., Chirenje, M., Stein, L., Abratt, R. & Wabinga, H. (2008). Part II: Cancer in Indigenous Africans—causes and control. The lancet oncology, 9, 786-795.

148. Stiller, C. A. (2004). Epidemiology and genetics of childhood cancer. Oncogene, 23, 6429.

149. Sulong, S. (2010). Genetic Aberrations in Childhood Acute Lymphoblastic Leukaemia: Application of High-Density Single Nucleotide Polymorphism Array. The Malaysian journal of medical sciences: MJMS, 17, 5.
150. Tasian, S. K. & Hunger, S. P. (2017). Genomic characterization of paediatric acute lymphoblastic leukaemia: an opportunity for precision medicine therapeutics. British journal of haematology, 176, 867-882.
151. Taspinar, M., Aydos, S. E., Comez, O., Elhan, A. H., Karabulut, H. G. & Sunguroglu, A. (2008). CYP1A1, GST gene polymorphisms and risk of chronic myeloid leukaemia. Swiss medical weekly, 138, 12-17.
152. Teachey, D. T. & Pui, C.-H. (2019). Comparative features and outcomes between paediatric T-cell and B-cell acute lymphoblastic leukaemia. The Lancet Oncology, 20, e142-e154.
153. Tedeschi, R., Bloigu, A., Ögmundsdottir, H. M., Marus, A., Dillner, J., Depaoli, P., Gudnadottir, M., Koskela, P., Pukkala, E. & Lehtinen, T. (2006). Activation of maternal Epstein-Barr virus infection and risk of acute leukemia in the offspring. American journal of epidemiology, 165, 134-137.
154. Tishkoff, S. A., Gonder, M. K., Henn, B. M., Mortensen, H., Knight, A., GIgnoux, C., Fernandopulle, N., Lema, G., Nyambo, T. B. & Ramakrishnan, U. (2007). History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. Molecular biology and evolution, 24, 2180-2195.
155. Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J. B., Awomoyi, A. A., Bodo, J.-M. & Doumbo, O. (2009). The genetic structure and history of Africans and African Americans. science, 324, 1035-1044.
156. Trang, N. T., Huyen, V. T., Tuan, N. T. & Phan, T. D. (2018). Association of N-acetyltransferase-2 and glutathione S-transferase polymorphisms with idiopathic male infertility in Vietnam male subjects. Chemico-biological interactions, 286, 11-16.
157. Uddin, M. M. N., Ahmed, M. U., Islam, M. S., Islam, M. S., Sayeed, M. S. B., Kabir, Y. & Hasnat, A. (2014). Genetic polymorphisms of GSTM1, GSTP1 and GSTT1 genes and lung cancer susceptibility in the Bangladeshi population. Asian Pacific Journal of Tropical Biomedicine, 4, 982-989.

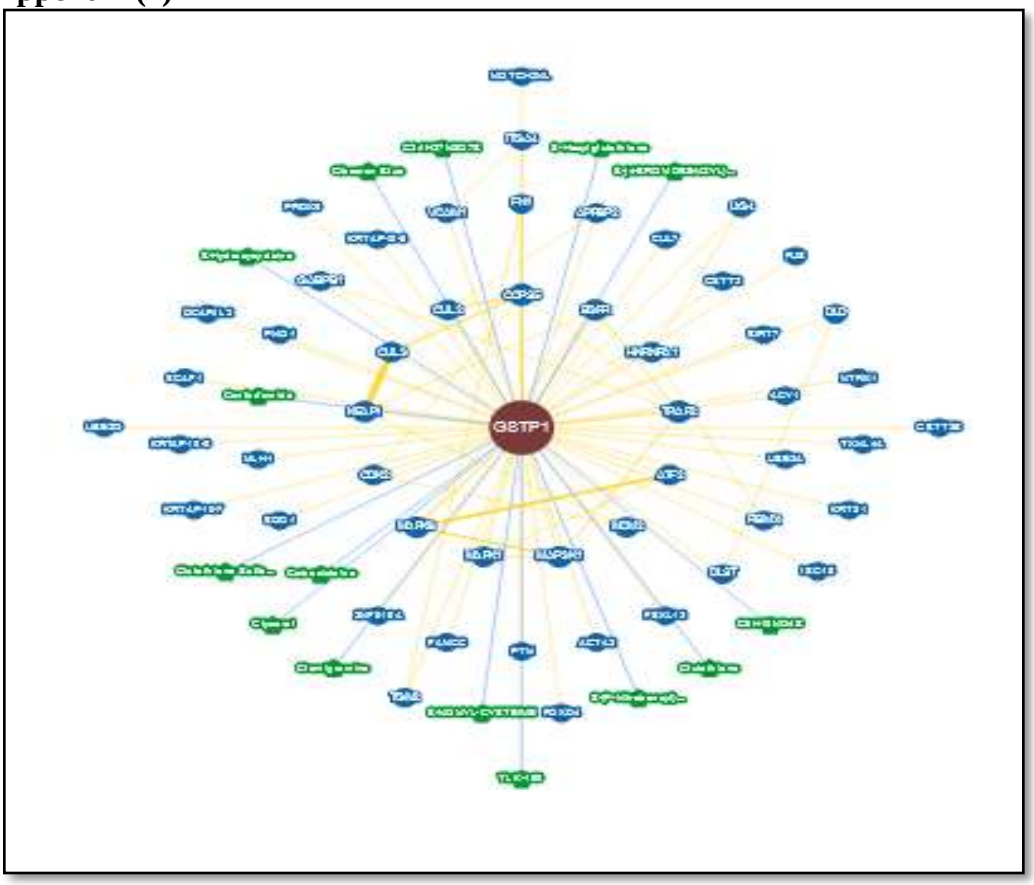
158. Vadillo, E., Dorantes-Acosta, E., Pelayo, R. & Schnoor, M. (2018). T cell acute lymphoblastic leukemia (T-ALL): new insights into the cellular origins and infiltration mechanisms common and unique among hematologic malignancies. Blood reviews, 32, 36-51.
159. Wenzinger, C., Williams, E. & Gru, A. A. (2018). Updates in the pathology of precursor lymphoid neoplasms in the revised fourth edition of the WHO classification of tumors of hematopoietic and lymphoid tissues. Current hematologic malignancy reports, 13, 275-288.
160. Williams, C. K. O. (2019). Risk Factors for Cancer. Cancer and AIDS. Springer.
161. Yabluchansky, M., Bogun, L., Martymianova, L., Bychkova, O., Lysenko, N. & Makienko, N. (2016). Clinical Symptoms of Blood System Affection: Anemia and Polycythemia; Leukopenias and Leukocytosis; Leukemia; Hemorrhagic Syndrome.
162. Yadav, D. S., Devi, T. R., Ihsan, R., Mishra, A. K., Kaushal, M., Chauhan, P. S., Bagadi, S. A., Sharma, J., Zamoawia, E. & Verma, Y. (2010). Polymorphisms of glutathione-S-transferase genes and the risk of aerodigestive tract cancers in the Northeast Indian population. Genetic testing and molecular biomarkers, 14, 715-723.
163. Yang, X., Qiu, M.-T., Hu, J.-W., Wang, X.-X., Jiang, F., Yin, R. & XU, L. 2013. GSTT1 null genotype contributes to lung cancer risk in asian populations: a meta-analysis of 23 studies. PloS one, 8, e62181.
164. Yu, Y., Li, X., Liang, C., Tang, J., Qin, Z., Wang, C., Xu, W., Hua, Y., Shao, P. & XU, T. 2016. The relationship between GSTA1, GSTM1, GSTP1, and GSTT1 genetic polymorphisms and bladder cancer susceptibility: A meta-analysis. Medicine, 95.
165. Zapletal, O., Procházková, J., Dubec, V., Hofmanová, J., Kozubík, A. & Vondráček, J. 2019. Butyrate interacts with benzo [a] pyrene to alter expression and activities of xenobiotic metabolizing enzymes involved in metabolism of carcinogens within colon epithelial cell models. Toxicology, 412, 1-11.
166. Zheng, Z., Qiao, Z., Gong, R., Yang, B., Xu, L., Zhang, L., Lu, Y., Jiang, B., Li, G. & Dong, C. 2017. Association between single nucleotide polymorphism of N-acetyltransferase 2 and susceptibility to acute lymphobl astic leukemia in

Chinese Han children. International journal of clinical and experimental pathology, 10, 7052-7058.

167. Zupa, A., Sgambato, A., Bianchino, G., Improta, G., Grieco, V., La torre, G., Campisi, B., Traficante, A., Aieta, M. & Cittadini, A. 2009. GSTM1 and NAT2 polymorphisms and colon, lung and bladder cancer risk: a case-control study. Anticancer research, 29, 1709-1714.

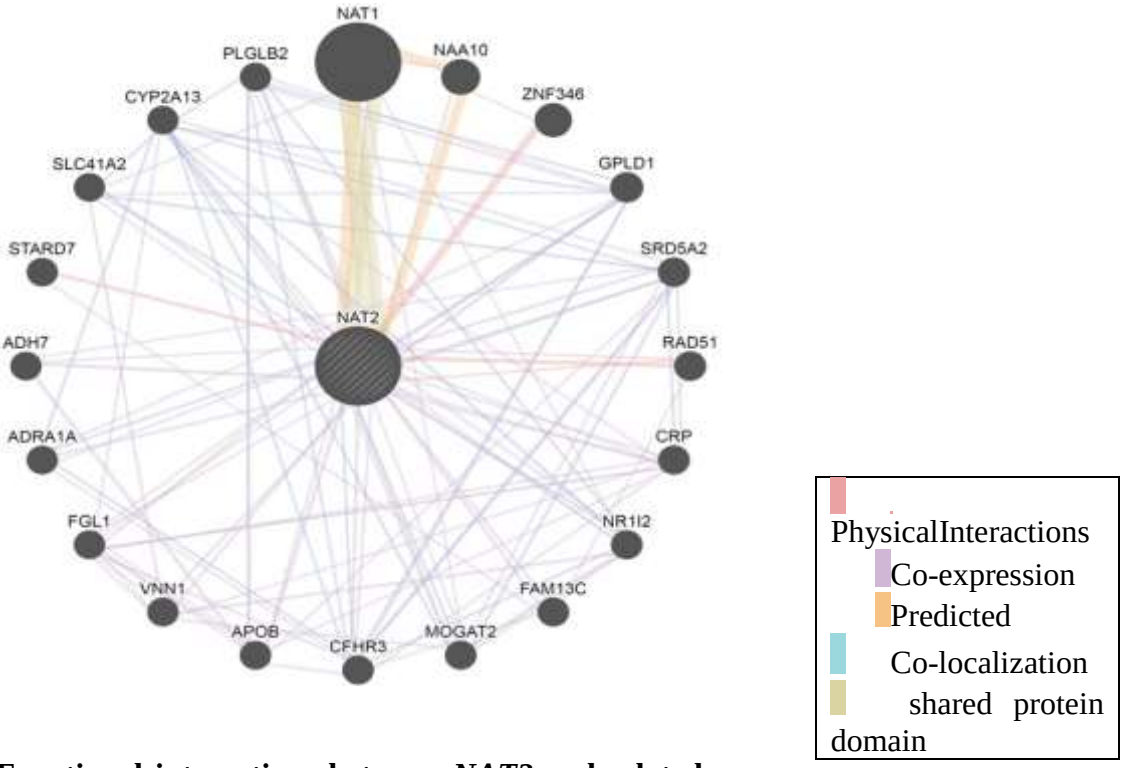
APPENDICES

Appendix (1)



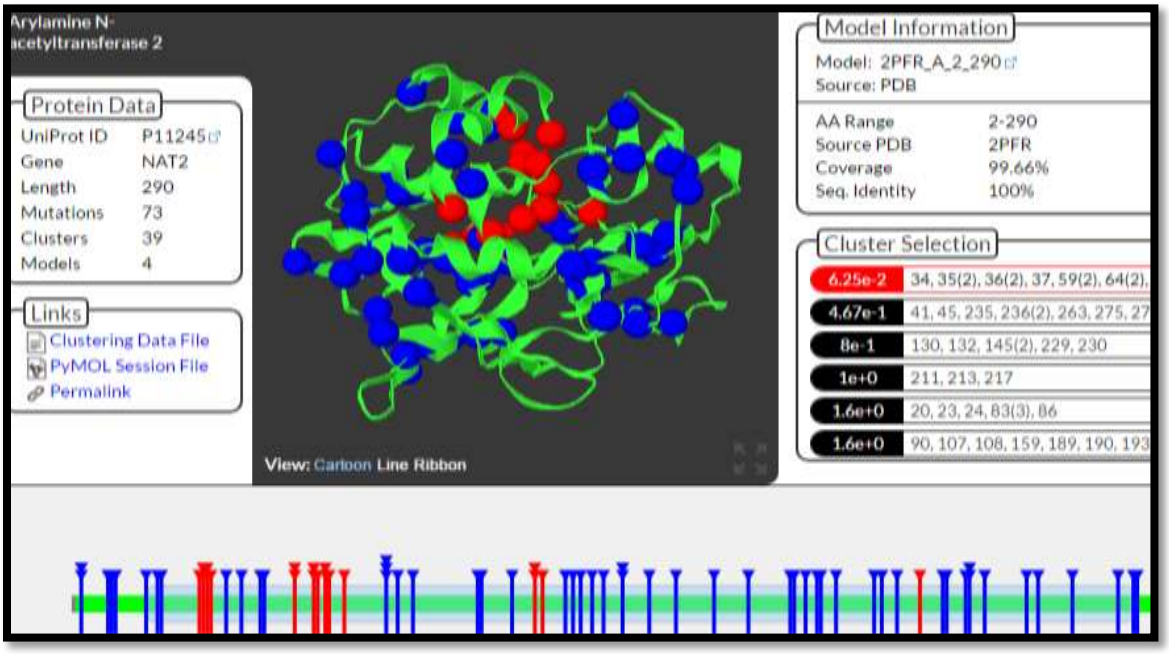
BioGRID result showGSTP1 interaction within other genes

Appendix (2)



Functional interactions between *NAT2* and related gene susing GeneMANIA software.

Appendix (3)



Distribution of extremely damaging mutations SNPs in NAT2 protein domains using Mutations3D

.